PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) **International Patent Classification:**
C12N 15/52, C12N 5/10, C12N 9/00, C12N 15/63, C12P 17/06

A2

(11) **International Publication Number:** **WO 00/22139**
(43) **International Publication Date:** 20 April 2000 (20.04.2000)

(21) **International Application Number:** PCT/US99/23535

(22) **International Filing Date:** 11 October 1999 (11.10.1999)

(30) **Priority Data:**
198 46 493.2 09 October 1998 (09.10.1998) DE

(60) **Parent Application or Grant**
GESELLSCHAFT FÜR BIOTECHNOLOGISCHE FORSCHUNG MBH (GBF) [/]; (). BRISTOL-MYERS SQUIBB COMPANY [/]; (). BEYER, Stefan [/]; (). BLOECKER, Helmut [/]; (). BRANDT, Petra [/]; (). CINO, Paul, M. [/]; (). DOUGHERTY, Brian, A. [/]; (). GOLDBERG, Steven, L. [/]; (). HOFLE, Gerhard [/]; (). MUELLER, Rolf-Joachim [/]; (). REICHENBACH, Hans [/]; (). BEYER, Stefan [/]; (). BLOECKER, Helmut [/]; (). BRANDT, Petra [/]; (). CINO, Paul, M. [/]; (). DOUGHERTY, Brian, A. [/]; (). GOLDBERG, Steven, L. [/]; (). HOFLE, Gerhard [/]; (). MUELLER, Rolf-Joachim [/]; (). REICHENBACH, Hans [/]; (). SANTUCCI, Ronald, R. ; ().

Published

(54) **Title: DNA SEQUENCES FOR ENZYMATIC SYNTHESIS OF POLYKETIDE OR HETEROPOLYKETIDE COMPOUNDS**
(54) **Titre: SEQUENCES D'ADN DESTINEES A LA SYNTHESE ENZYMATIQUE DE COMPOSES A BASE DE POLYKETIDES OU D'HETEROPOLYKETIDES**

(57) Abstract

The invention consists of: (1) cloned Sorangium cellulosum polyketide synthase (PKS) biosynthetic cluster DNA; and (2) the nucleotide sequence and predicted protein coding sequences of the cloned DNA. The invention can be used for, but not limited to: (a) increasing yields of PKS product in Sorangium cellulosum (e.g., by amplification or genetic modification of the epothilon gene cluster or its component parts); (b) increasing yields of polyketide product in a heterologous system by transfer of the epothilone gene cluster or its component parts, which may be followed by amplification or genetic modification of the PKS gene cluster or its component parts; (c) modification of the polyketide product chemical structure in either Sorangium cellulosum or a heterologous host (e.g., by genetic modification of the epothilone gene cluster or its component parts; and (d) for the detection of genes and gene products involved in making polyketides or related molecules in other organisms (e.g., by hybridization or complementation assays). DNA sequence and analysis is presented for the following cosmids and plasmids: A2 cosmid; the pEPOcos6 region (overlapping of pEPOcos6 and pEPOcos7); pEPOcos8 cosmid; A5 cosmid; Sau4 (10 kb plasmid).

(57) Abrégé

L'invention concerne: (1) un ADN biosynthétique cloné en grappe de polykétide synthase (PKS) de Sorangium cellulosum; et (2) la séquence nucléotidique et les séquences protéiques codantes prévues de l'ADN cloné. L'invention peut avoir les applications suivantes (sans caractère limitatif): (a) augmentation de la production de PKS chez Sorangium cellulosum (p.ex., par l'amplification ou la modification génétique de la grappe de gènes épothilone ou de ces parties constitutives); (b) augmentation de la production du produit polykétide dans un système hétérologue par le transfert de la grappe de gènes épothilone ou de ces parties constitutives, qui peut être suivie par l'amplification ou la modification génétique de la grappe de gènes PKS ou des ses parties constitutives; (c) modification de la structure chimique du produit polykétide soit chez Sorangium cellulosum soit chez un hôte hétérologue (p.ex., par l'amplification ou la modification génétique de la grappe de gènes épothilone ou de ces parties constitutives); et (d) détection de gènes et de produits géniques participant à la fabrication de polykétides ou de molécules correspondantes dans d'autres organismes (p.ex., par des dosages à hybridation ou à complémentation). La séquence d'ADN et l'analyse sont présentées pour les cosmides et les plasmides suivants: cosmide A2; région pEPOcos6 (se chevauchant avec pEPOcos6 et pEPOcos7); cosmide pEPOcos8; cosmide A5; Sau4 (plasmide 10 kb).

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/52, 9/00, 15/63, 5/10, C12P 17/06	A2	**(11) International Publication Number:** WO 00/22139 **(43) International Publication Date:** 20 April 2000 (20.04.00)

(21) International Application Number: PCT/US99/23535

(22) International Filing Date: 11 October 1999 (11.10.99)

(30) Priority Data:
198 46 493.2 9 October 1998 (09.10.98) DE

(71) Applicants *(for all designated States except US)*: GESELLSCHAFT FÜR BIOTECHNOLOGISCHE FORSCHUNG MBH (GBF) [DE/DE]; Mascheroder Weg 1, D-38124 Braunschweig (DE). BRISTOL-MYERS SQUIBB COMPANY [US/US]; P.O. Box 4000, Princeton, NJ 08543-4000 (US).

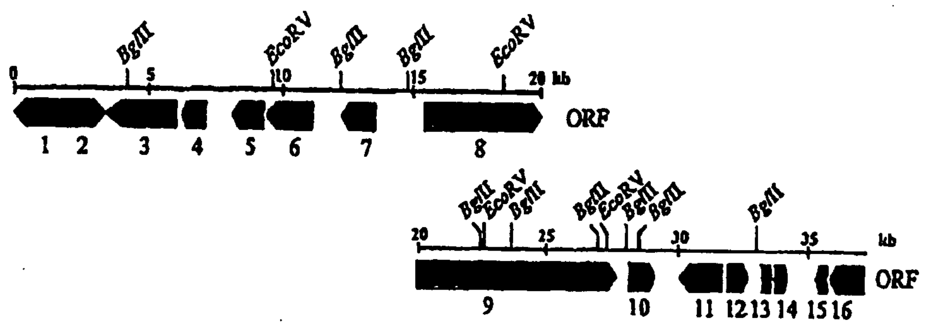
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BEYER, Stefan [DE/DE]; Mascheroder Weg 1, D-38124 Braunschweig (DE). BLOECKER, Helmut [DE/DE]; Mascheroder Weg 1, D-38124 Braunschweig (DE). BRANDT, Petra [DE/DE]; Mascheroder Weg 1, D-38124 Braunschweig (DE). CINO, Paul, M. [US/US]; 4 Crest Drive, Bound Brook, NJ 08805 (US). DOUGHERTY, Brian, A. [US/US]; 10 Rosemary Lane, Killingworth, CT 06419 (US). GOLDBERG, Steven, L. [US/US]; 25 Compton Court, Basking Ridge, NJ 07920 (US). HOFLE, Gerhard [DE/DE]; Mascheroder Weg 1, D-38124 Braunschweig (DE). MUELLER, Rolf-Joachim [DE/DE]; Mascheroder Weg 1, D-38124 Braunschweig (DE). REICHENBACH, Hans [DE/DE]; Mascheroder Weg 1, D-38124 Braunschweig (DE).

(74) Agents: SANTUCCI, Ronald, R. et al.; Kane, Dalsimer, Sullivan and Levy, LLP 20th floor, 711 Third Avenue, New York, NY 10017 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: DNA SEQUENCES FOR ENZYMATIC SYNTHESIS OF POLYKETIDE OR HETEROPOLYKETIDE COMPOUNDS



(57) Abstract

The invention consists of: (1) cloned *Sorangium cellulosum* polyketide synthase (PKS) biosynthetic cluster DNA; and (2) the nucleotide sequence and predicted protein coding sequences of the cloned DNA. The invention can be used for, but not limited to: (a) increasing yields of PKS product in *Sorangium cellulosum* (e.g., by amplification or genetic modification of the epothilone gene cluster or its component parts); (b) increasing yields of polyketide product in a heterologous system by transfer of the epothilone gene cluster or its component parts, which may be followed by amplification or genetic modification of the PKS gene cluster or its component parts; (c) modification of the polyketide product chemical structure in either *Sorangium cellulosum* or a heterologous host (e.g., by genetic modification of the epothilone gene cluster or its component parts; and (d) for the detection of genes and gene products involved in making polyketides or related molecules in other organisms (e.g., by hybridization or complementation assays). DNA sequence and analysis is presented for the following cosmids and plasmids: A2 cosmid; the pEPOcos6 region (overlapping of pEPOcos6 and pEPOcos7); pEPOcos8 cosmid; A5 cosmid; Sau4 (10 kb plasmid).

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

Description

DNA sequences for enzymatic synthesis of polyketide or heteropolyketide compounds

The present invention relates to DNA sequences for enzymatic synthesis of polyketide or heteropolyketide compounds produced by the bacterium *Sorangium cellulosum*.

Background and introduction

This patent application describes DNA sequences for the enzymatic synthesis of polyketide and/or heteropolyketide structures synthesized by the myxobacterium *Sorangium cellulosum*. Several of these compounds have known cytotoxic, immunosuppressive, antibiotic and fungicidal biological activity, with the epothilones having been most studied and characterized. The fermentation of large quantities of secondary metabolites from microorganisms, especially from myxobacteria, is a time consuming and difficult process that often involves complications (e.g. contamination, low product yield, difficult isolation and purification). Therefore it would be advantageous to use a well-characterized organism for such fermentations. After cloning of the desired biosynthetic genes one could create such an organism via genetic engineering and manipulate the biosynthesis of the compound. Identified sequences can be cloned into optimized expression vectors and generate recombinant cell lines that overproduce polyketide structures.

Polyketide synthases (PKS) and non-ribosomal peptide synthetases (NRPS) represent macromolecular and multifunctional enzymes which are characterized by a modular architecture. PKS condenses activated carbonic acids (usually acetate and propionate) and reduce the resulting 2-keto acid intermediates stepwise in a fatty acid biosynthesis-like fashion. Responsible for each reaction step is a specific domain that recognizes, activates, condenses and reduces the carbonic acid. Depending on the presence of these domains in the corresponding modules, every reduction stage can occur in the final product (Rawlings, Nat. Prod. Reports <u>14</u>, 523-556 [1997]; for a review, see Chem. Rev. <u>97</u>, 2463-2760 [1997]). A typical example for the biosynthesis of a polyketide is the macrolide antibiotic erythromycin (Staunton and Wilkinson, Chem. Rev. <u>97</u>, 2611-2630 [1997]). NRPSs are also modular enzymes and condense via peptide bonds amino acids to low molecular weight bioactive substances like bacitracin or tyrocidin. Typical domains of these systems activate the amino acid and condense it with the growing peptide chain. Methylations, epimerisations and modifications via additional protein domains are possible (Stachelhaus and Marahiel, FEMS Microbiol Lett. <u>125</u>, 3-14 [1995]). Both types of enzymes (NRPS and PKS) share the modular organization of the proteins in which specific catalytic domains are responsible for recognition, activation, condensation and modification of the single elongation units. The growing chain of amino acids and/or carbonic acids is extended through the action of one module adding one unit. The domains of each module carry the active centers responsible for the enzymatic steps of the biosynthesis.

Little is known about the biosynthesis of biologically active polyketides and polypeptides from myxobacteria. Fragments of the biosynthetic gene clusters of soraphen and saframycin have been described from *Sorangium cellulosum* So ce26 and *Myxococcus xanthus*, respectively (Schupp et al., J. Bacteriol. 177, 3673-3679 [1995] and Pospiech et al., Microbiology 141, 1793-1803 [1995]). We have constructed genomic libraries of the epothilone producer *Sorangium cellulosum* So ce90. Gene probes based on PKS and PS genes were used to isolate recombinant cosmids, which were then sequenced and characterized. Several unique pathways containing PKS, PS, or a combination of both types of genes were identified, demonstrating that this organism is potentially a rich source of novel bioactive compounds.

A subject of the present invention is therefore to provide DNA sequences according to claim 1 the expression products of which perform or are involved in the enzymatic biosynthesis, mutasynthesis or partial synthesis of polyketide or heteropolyketide compounds. The DNA sequences may be inserted into well known and optimized expression vectors by commmon techniques of molecular biology, thus allowing transformation, selection and cloning of cells, which cells are then capable of synthezising polyketide or heteropolyketide compounds by fermentation. Using an overproducing clone allows the desired polyketide or heteropolyketide compounds be easily produced and recovered in high amounts. Further, knowledge of the localization of regulatory DNA segments and individual structural genes allows "site-directed mutagenesis" using common techniques for genetic engineering, and thus construction of optimized enzymes ("protein engineering") for fermentative synthesis of polyketide or heteropolyketide compounds.

The invention thus further relates to a recombinant expression vector according to claim 16, cells transformed therewith according to claim 17 and to a process for enzymatic biosynthesis, mutasynthesis or partial synthesis of polyketide or heteropolyketide compounds according to claim 23.

Preferred and/or advantageous embodiments of the present invention are subject-matter of the subclaims.

In brief, the invention consists of (1) cloned *Sorangium cellulosum* polyketide synthase (PKS) and/or peptide synthetase (PS) biosynthetic cluster DNA and (2) the nucleotide sequence and predicted protein coding sequences of the cloned DNA. The invention can be used for, but not limited to, (a) increasing yields of PKS product in *Sorangium cellulosum* (e.g., by amplification or genetic modification of the epothilone gene cluster or its component parts), (b) increasing yields of polyketide and/or peptide synthetase product in a heterologous system by transfer of the corresponding gene cluster or its component parts, which may be followed by amplification or genetic modification of the PKS and/or PS gene cluster or its component parts, (c) modification of the polyketide and/or peptide synthetase product chemical structure in either *Sorangium cellulosum* or a heterologous host (e.g., by genetic modification of the corresponding gene cluster or its component parts) and (d) for the detection of genes and gene products involved in making polyketides or related molecules in other organisms (e.g., by hybridization or complementation assays). DNA sequence and analysis is presented for the following cosmids and plasmids:

- A2 cosmid as defined in claim 6
- the pEPOcos6 region (overlapping of pEPOcos6 and pEPOcos7) as defined in claim 7

- pEPOcos8 cosmid as defined in claim 10
- A5 cosmid as defined in claim 12
- Sau4 (10 kb plasmid) as defined in claim 14

The invention is now described in more detail by examples and for illustration only. The examples are not to be construed as any limitation of the scope.

Figure 1 is a restriction map of one of the DNA sequences of the present invention (cosmid A2 insert) indicating also the localization of regulatory DNA segments and the individual structural genes ("open reading frames" or ORFs) 1 to 16.

Figure 2 shows the open reading frames found on pEPOcos6 region

DNA sequence data from A2 cosmid are as defined in claim 6.

Table 1 correlates ORFs 1 to 16 found on A2 cosmid with the respective biological function (Regulators, Enzymes).

Table 1

	gene/function	position
ORF 1	regulatory element	1666 - 1
ORF 2	regulatory element	1605 - 3338
ORF 3	acyl-t-RNA synthetase	6100 - 3398
ORF 4	monooxygenase	7110 - 6374
ORF 5	amino transferase	9590 - 8433
ORF 6	L-dopa decarboxylase	11393 - 9855
ORF 7	oxidoreductase	13656 - 12712
ORF 8	polyketide synthase	15374 - 18984
ORF 9	polypeptide synthetase	20003 - 27889
ORF 10	peptidase	28251 - 29402
ORF 11	regulatory element	31720 - 30401
ORF 12	sigma factor	31982 - 32932
ORF 13	regulatory element	33128 - 33613
ORF 14	regulatory element	33661 - 34007
ORF 15	transcription regulator	35611 - 35255
ORF 16	signal transduction	37856 - 35730

Working Examples

A. Construction of a *Sorangium cellulosum* cosmid library

1. Isolation of genomic DNA from *S. cellulosum* So ce90

a. *Sorangium cellulosum* So ce90 was spread onto solid CA-2 agar and incubated at 30°C for 5-7 days. CA-2 agar is prepared by autoclaving 18 g Bacto-agar (Difco Laboratories, Detroit, MI) in 800 ml dH_2O for 20 min at 121°C and cooling to 50-55°C in a water bath. The following filter-sterilized solutions are added to the agar: 20% (w/v) glucose, 50 ml; Solution A (7.5% [w/v] KNO_3, 7.5% K_2HPO_4), 10 ml; Solution B (1.5% [w/v] $MgSO_4 \cdot 7H_2O$), 10 ml; Solution C (0.2% [w/v] $CaCl_2 \cdot 2H_2O$, 0.15% [w/v] $FeCl_3$), 10 ml; 1 M HCl, 1 ml; autoclaved 4-day old *Sorangium cellulosum* broth, 100 ml. A sample of cells was removed from the plates with a sterile loop and inoculated into 50 ml of G51t medium in a 250 ml Erlenmeyer flask. G51t consists of 0.5% starch (Cerestar), 0.2% tryptone, 0.1% yeast extract, 0.05% $CaCl_2$, 0.05% $MgSO_4 \cdot 7H_2O$, 1.2% 4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid (HEPES), 0.2% glucose, pH 7.6. The flasks were shaken at 30°C, 160 rpm until a dense orange bacterial growth was obtained (ca. 5-7 d.). The cells were pelleted by centrifugation at 6,000 x g and used immediately or stored frozen at -20°C.

The protocol used for isolating chromosomal DNA from bacteria using hexadecyltrimethylammmonium bromide (CTAB) has been described previously (Ausubel et al., Current Protocols in Molecular Biology, John Wiley and Sons, New York, 1990). The precipitated DNA was recovered with a bent Pasteur pipette, washed with 70% and 95% ethanol, air-dried, and resuspended in 0.5 ml TE buffer (0.01 M Tris-HCl, 0.001 M ethylenediaminotetraacetic acid [EDTA], pH 8.0).

b. Alternatively, genomic DNA was isolated from *S. cellulosum* cells cultured as described in section A.1 using the Midi Qiagen Blood & Cell Culture DNA purification Kit (Qiagen, Hilden, Germany) following the Qiagen Genomic DNA Handbook protocol for bacterial DNA isolation (1997, Qiagen, Hilden, Germany, p. 29 ff.). In order to obtain high molecular weight chromosomal DNA the precipitated DNA was recovered with a bent pasteur pipette as described in section A.1.

2. Isolation of plasmid DNA

a. pFD666: pFD666 is a bifunctional *E. coli-Streptomyces* cosmid cloning vector (see Denis and Brzezinski, Gene 111, 115-118 [1992]). To maintain stability of large inserts, it is present in low-medium copy number when replicated in *E. coli*. For this reason, isolation of sufficient pure DNA to carry out cloning experiments was difficult using commercial kits with standard protocols. A modified procedure was therefore used to obtain pFD666 DNA. A 10 ml culture of DH10B(pFD666) was grown for 16-20 hr at 37°C in LB (1% tryptone, 0.5% yeast extract, 0.5% NaCl, pH 7.0) medium containing 50 µg/ml kanamycin sulfate. Fifty ml of LB + kanamycin was inoculated to a starting OD_{600} of ca. 0.25 and shaken at 300 rpm, 37°C, until the OD_{600} reached ca. 0.6. Five hundred ml of LB + kanamycin medium in a 2 l flask was inoculated with 25 ml of this culture and incubated under the same conditions for 2.5 hr. Chloramphenicol (2.5 ml of a 34 mg/ml solution in 100% EtOH) was added and the incubation continued for an additional 16-20 hr. (The previous steps were performed according to Maniatis et al. Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989.) Cells were pelleted for 10 min, 16,000 x g . They were resuspended in 9 ml of 50 mM glucose/25 mM Tris-HCl (pH 8.0)/10 mM EDTA and transferred to a 50 ml disposable centrifuge tube. One ml of a freshly-prepared 10 mg/ml lysozyme solution in 10 mM Tris-HCl, pH 8.0 was added and the cell suspension incubated in a 37°C water bath for 10 min. Twenty ml of a freshly-prepared 0.2 NaOH/1% sodium dodecyl sulfate (SDS) solution was added and the tube inverted gently 5-7 times to mix the contents. After 5 min at room temperature, 15 ml of 5 M potassium actate (pH 4.8) was added and the tube inverted sharply 3-4 times. The tube was centrifuged at 6,000 x g for 10 min at 4°C and the supernatant poured though 2 layers of sterile cheese cloth into a fresh 50 ml disposable tube. Isopropanol to a final concentration of 0.6% was added and the contents of the tube mixed several times. The precipitated nucleic acid was centrifuged at 6,000 x g for 10 min at 4°C. The pellet was washed with 70% EtOH and any excess EtOH was aspirated from the pellet, which was allowed to air dry for 5 min. It was resuspended in 5 ml of 50 mM 3-(N-Morpholino)propanesulfonic acid (MOPS)/750 mM NaCl, pH 7.0 and added to an equilibrated to QIAfilter Midi column (Qiagen, Chatsworth, CA). The manufacturer's protocol for washing and eluting the plasmid DNA was followed.

b. **SuperCos:** SuperCos plasmid DNA was purchased from Stratagene (La Jolla, CA).

3. Preparation of ca. 38-47 kb *Sau*3A1 fragments of *S. cellulosum* chromosomal DNA

a. *S. cellulosum* chromosomal DNA prepared as described in section A.1.a was partially cleaved with restriction endonuclease *Sau*3A1 in a 1000 μl reaction volume consisting of 50 μg chromosomal DNA, 5 units enzyme (Promega, Madison , WI), 0.006 M Tris-HCl, 0.006 M $MgCl_2$, 0.10 M NaCl, and 0.001 M dithiothreitol (pH 7.5) for 5 min at 37°C. The reaction mixture was extracted once with an equal volume of 1:1 phenol:chloroform. After centrifugation, the upper aqueous phase was saved, to which 0.1 vol. of 3 M sodium acetate and 0.6 vol. isopropanol was added. DNA was pelleted by centrifugation for 5 min at 16,000 x g in a microfuge and washed once with 0.5 ml 70% EtOH. After drying in a SpeedVac (Savant Instruments, Farmingdale, NY) for 5 min, the pellet was resuspended in 0.1 ml TE buffer. The DNA was layered ontop of a 12 ml 10-40% sucrose gradient prepared in TE buffer and centrifuged at 113,600 x g for 16 hr, 10°C using a Beckman SW40Ti rotor (Beckman Instruments, Palo Alto, CA). Five hundred μl aliquots of the gradient were removed using a pipetor beginning at the top of the tube. Samples (5 μl) of the fractions were analyzed by electrophoresis through a 0.5% agarose gel in TAE buffer (0.04 M Trizma base, 0.02 M acetic acid, and 0.001 M EDTA, pH 8.3) containing 0.5 μg/ml ethidium bromide for 6 hr at 100 V. Fractions containing DNA fragments of ca. 40-45 kb were identified by comparison to a high molecular weight DNA standard (Life Technologies, Gaithersburg, MD). Sucrose was diluted from the corresponding 0.5 ml fraction by addition of 0.5 vol. TE. Subsequently, DNA was precipitated by addition of 0.1 vol. 3 M sodium acetate and 0.6 vol. isopropanol. DNA was pelleted by centrifugation at 16,000 x g for 10 min in a microfuge. DNA was washed with 0.5 ml 70% EtOH and dried in a SpeedVac with moderate heat for 10 min. Finally, the DNA was resuspended in distilled H_2O at a concentration of 0.5 mg/ml.

b. Alternatively, 10 µg of *S. cellulosum* chromosomal DNA prepared as described in A.1.b was treated with 0.3 U *Sau*3A1 (New England Biolabs, Beverly, MA) for 1 h at 37°C in 400 µl of the supplier's recommended reaction buffer. Formation of DNA fragments of about 40 kb in size was checked by comparison of the motility behavior with high molecular weight DNA standards after a 0.3% agarose gel electrophoresis. An equal volume of phenol:chloroform (1:1) was added, mixed and centrifuged. The upper aqueous phase was recovered and 0.1 vol. of 3 M sodium acetate and 0.6 vol. of isopropanol were added. After centrifugation, the precipitated DNA was washed twice with 0.5 ml 70% ice cold ethanol and finally air-dried. The DNA fragments were resuspended in 100 µl shrimp alkaline phosphatase reaction buffer and dephosphorylated for 150 min. at 37°C using 2 U shrimp alkaline phosphatase (Amersham Life Science, Cleveland, OH). A phenol:chloroform extraction followed as described above. Finally, the DNA was precipitated by addition of 0.1 vol. 3 M sodium acetate and 0.6 vol. isopropanol, dried, and dissolved in TE buffer.

4. Preparation of cosmid libraries

a. Using pFD666: Vector pFD666 was cleaved with restriction endonuclease BamHI in a 0.02 ml reaction volume consisting of 2 μg plasmid DNA, 10 units of BamHI (Promega), 0.006 M Tris-HCl, 0.006 M $MgCl_2$, 0.05 M NaCl, and 0.001 M dithiothreitol (pH 7.5) for 90 min at 37°C. Five μl of 10x alkaline phosphatase buffer (0.5 M Tris-HCl [pH 9.3], 0.01 M $MgCl_2$, 0.001 M $ZnCl_2$, 0.01 M spermidine) was added to the reaction followed by alkaline phosphatase (0.01 units/pmol ends; Promega) and distilled H_2O to a final volume of 0.05 ml. The sample was incubated for 30 min at 37°C and a second aliquot of phosphatase was added. After a further 30 min at 37°C, 0.3 ml of stop buffer (0.01 M Tris-HCl [pH 7.5], 0.001 M EDTA, 0.2 M NaCl, 0.5% SDS) and 0.35 ml of 1:1 phenol; $CHCl_3$ was added to the reaction. The sample was mixed gently several times by inversion and centrifuged at 16,000 x g for 3 min to separate the phases. The aqueous layer was removed to a new microfuge tube. 0.1 vol. 3 M sodium acetate and 2 vol. 100% EtOH were added and the precipitated DNA pelleted by centrifugation at 16,000 x g for 10 min. Liquid was removed by aspiration and the pellet washed once with 0.5 ml 70% EtOH. The DNA was dried in a SpeedVac and resuspended in TE buffer to 0.5 mg/ml.

Digested, phosphatase-treated pFD666 was ligated to the partially-cleaved chromosomal DNA (see sections A.3.a and B.1.a) in a 0.005 ml reaction consisting of 1 μg pFD666, 1 μg *S. cellulosum* DNA, 0.03 M Tris-HCl (pH 7.8), 0.01 M $MgCl_2$, 0.01 M dithiothreitol, and 0.0005 M adenosine-5′-triphosphate and 1.5 Weiss units of T4 DNA ligase (Promega). The reaction was carried out at room temperature for 2 hr. The entire reaction mix was packaged into bacteriophage λ *in vitro* using Packagene extracts (Promega) according to the manufacturer's directions. The entire packaging reaction (0.5 ml) was diluted with 4.5 ml SM buffer (per liter: 5.8 g NaCl, 2 g $MgSO_4.7H_20$, 1 M Tris-HCl[pH 7.5], 5 ml 2% gelatin solution). Transfection was performed by adding 10 ml of an overnight culture of *E. coli* DH5α that had been grown in LB medium with 0.01 M $MgSO_4$ and 0.2% maltose to the diluted phage and incubating at 37°C for 20 min. 0.8 ml of LB was added and the cells shaken at 225 rpm for 1 hr at 37°C. Cells were pelleted, resuspended in LB, and spread onto a 150 mm LB + kanamycin agar plate. After 3 d. at 30°C, the colonies were harvested by picking ca. 800 colonies into 2.0 ml LB + kanamycin medium containing 20% glycerol, freezing on dry ice, and storing at -70°C. In addition, six kanamycin-resistant colonies were inoculated into 2 ml LB + kanamycin liquid medium and incubated at 37°C, 250 rpm, for 18-24 hr. Cosmid DNA was prepared using a standard alkaline lysis procedure starting with 1.5 ml of the culture. DNA was digested with restriction endonuclease PstI and samples electrophoresed on a 0.8% TAE agarose gel for 1.5 hr at 100 V. A unique restriction pattern was noted in each sample and the total size of the insert was calculated to be between 40 and 45 kilobases.

b. Using SuperCos: 30 μg of vector SuperCos was digested with XbaI (New England Biolabs, Beverly, MA) for 210 min at 37°C in 100 μl of the recommended reaction buffer. Ten μl sodium acetate and 60 μl isopropanol was added before the solution was centrifuged for 30 min at 16,000 x g. The precipitated DNA was washed twice with 500 μl ice cold 70% ethanol. The vector DNA was precipitated and air-dried, dissolved in 135 μl shrimp alkaline phosphatase reaction buffer and treated with 2.5 U shrimp alkaline phosphatase for 150 min. After heat inactivation of the enzyme at 75°C for 20 min, a phenol:chloroform extraction was performed as described in section 1. c. The DNA, resuspended in 100 μl BamHI restriction buffer was hydrolyzed with 15 U BamHI (New England Biolabs, Beverly, MA) for 180 min. A phenol:chloroform extraction followed (see section A.3). The SuperCos DNA was precipitated by additon of 0.1 vol 3 M sodium acetate and 0.6 vol isopropanol, centrifuged at 16,000 x g, and resuspended in 50 μl TE buffer.

Four μg of digested vector DNA was ligated with 10 μg partially hydrolyzed genomic DNA from *S. cellulosum* (as described in section A.3.b) in a final volume of 20 μl using 2 U T4 DNA ligase and the appropriate reaction buffer (Gibco BRL, Eggenstein, Germany). The reaction was carried out at 16°C overnight. The reaction mixture was packaged into phage particles using the Gigapack III XL packaging extract kit (Stratagene) according to the manufacture´s protocol. Treatment of packaging reaction mixture and transfection of *E. coli* SURE (Stratagene) was performed as described in 4.a. Transfected cells were concentrated by centrifugation, resuspended in fresh LB medium and distributed on LB agar plates containing 50 $\mu g/ml^{-1}$kanamycin. The plates were incubated overnight at 30°C. 1600 recombinant clones were transferred into 96 well microtiter plates filled with 80 μl LB medium containing 50 μg/ml˙kanamycin per well and propagated overnight at 30°C. The following day the microtiter plates were used to inoculate a second set of microtiter plates in order to obtain a duplicate of the recombinant clones. Each well of the original set of microtiter plates was supplemented with 80 μl 50 % glycerol and the entire plate stored at -70°C.

20 randomly chosen transformants were inoculated into 3 ml LB medium with 50 $\mu g/ml^{-1}$ kanamycin and incubated over night at 37°C in order to isolate plasmid DNA using the Qiagen plasmid extraction kit (Qiagen, Hilden, Germany). Restriction fragment analysis of the recombinant cosmids using the restriction endonucleases PstI and BglII indicated that the cosmids contained inserts of approximately 35 to 42 kb in size.

B. Construction of a *S. cellulosum* plasmid library

1. Preparation of 8-12 kb fragments of *S. cellulosum* chromosomal DNA.

S. cellulosum chromosomal DNA prepared as described in section A.1.a was partially cleaved with restriction endonuclease Sau3A1 in a 100 µL reaction volume consisting of 5 µg chromosomal DNA, 5 units enzyme (Promega, Madison , WI), 0.006 M Tris-HCl, 0.006 M $MgCl_2$, 0.10 M NaCl, and 0.001 M dithiothreitol (pH 7.5) for 4 min at 37°C. The digested DNA was electrophoresed through a 11 x 14 cm 0.8% TAE-agarose gel for 18 hr at 17 V. Fragments of 8-12 kb were cut from the gel and purified using the QIAquick Gel Extraction Kit using the manufacturer's protocol (Qiagen).

2. Preparation of the plasmid library

Plasmid pZero2.1 (Invitrogen, Carlsbad, CA) was cleaved with restriction endonuclease BamHI in a 0.02 ml reaction volume consisting of 1 µg plasmid DNA, 10 units of BamHI (Promega), 0.006 M Tris-HCl, 0.006 M $MgCl_2$, 0.05 M NaCl, and 0.001 M dithiothreitol (pH 7.5) for 20 min at 37°C. 0.08 ml of dH_2O and 0.1 ml of 1:1 phenol:CHCl3 was added. The sample was briefly vortexed and centrifuged at 16,000 x g for 2 min. The aqueous layer was removed to a new microfuge tube. 0.1 vol. 3 M sodium acetate and 2 vol. 100% EtOH were added and the precipitated DNA pelleted by centrifugation at 16,000 x g for 10 min. Liquid was removed by aspiration and the pellet washed once with 0.5 ml 70% EtOH. The DNA was dried in a SpeedVac and resuspended in TE buffer to 0.004 μg/ml. Digested pZero2.1 was ligated to the partially-cleaved chromosomal DNA in a 0.01 ml reaction consisting of 0.004 μg pZero2.1, 0.05 μg *S. cellulosum* DNA, 0.03 M Tris-HCl (pH 7.8), 0.01 M $MgCl_2$, 0.01 M dithiothreitol, and 0.0005 M adenosine-5'-triphosphate and 1.5 Weiss units of T4 DNA ligase (Promega). The reaction was carried out at room temperature for 2 hr. 0.015 ml dH_2O and 0.25 ml of 1-butanol were added, the sample vortexed briefly, and centrifuged at 16,000 x g for 10 min. Liquid was aspirated away from the pellet and the sample dried in a SpeedVac for 5 min. The ligated DNA was resuspended in 0.005 ml dH_2O and mixed with 0.04 ml of electrocompetent *Escherichia coli* DH10B cells (GIBCO/BRL, Gaithersburg, MD). The sample was placed into a pre-chilled 0.2 mm-gap electroporation cuvette and transformed into the bacteria by electroporation using a BioRad Gene Pulser II unit (BioRad, Hercules, CA) at 25 μF and 200 Ω. 0.96 ml SOC medium (0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM $MgSO_4$, 20 mM glucose) was mixed with the cells and transferred to a 1.5 ml microfuge tube. The sample was incubated at 37°C, 225 rpm, for 1 hr. Aliquots of the cells were spread onto an LB agar + kanamycin and incubated at 37°C for 20 hr to estimate the number of transformants obtained. Six kanamycin resistant colonies were confirmed to contain an insert of the expected size as described in section A.4.a.

C. Identification of cosmids possessing polyketide synthase genes

1. Colony blot hybridizations using cosmid library in pFD666: A 20 x 20 cm sheet of Duralon UV membrane (Stratagene) was placed on top of a 24.5 x 24.5 cm square bioassay dish containing 250 ml LB agar - kanamycin. An aliquot of the frozen cosmid library in 1 ml LB medium was spread on the filter. The plate was incubated at 37°C for 24 hr. Colonies were replicated onto two fresh filters which were placed onto LB + kanamycin agar medium and incubated at 28°C for 18 hr. Lysis of cells and neutralization of released DNA was performed according to directions that were provided with the filters. The DNA was crosslinked to the filters using a UV Stratalinker 2400 unit (Stratagene) in the auto crosslink mode. Cell debris was removed by placing the filters in a container with a solution of 3 X SSC (20 X SSC contains, per liter, 173.5 g NaCl, 88.2 g sodium citrate, pH adjusted to 7.0 with 10 N NaOH), 0.1% SDS and rubbing the lysed colonies with a Kimwipe. The filters were then incubated at least 3 hr with the same wash solution for at least 3 hr at 65°C. The plasmid library was treated similarly except cells were spread onto a 137 mm circular Duralon UV membrane placed on top of a 150 mm petri dish containing 80 ml LB agar + kanamycin.

For hybridizations, a probe consisting of a 650-base pair (bp) polymerase chain (PCR) fragment representing a portion of a *S. cellulosum* polyketide synthase gene was used. The fragment was amplified using primers to consensus regions of Type I (macrolide) polyketide synthase (PKS) genes (Swan et al., Mol. Gen. Genetics <u>242</u>, 358-362 [1994]). A series of sense and antisense oligonucleotides were prepared for PCR studies as indicated in the following table 2:

Table 2

Oligo-nucleotide	<u>I. DNA sequence (5'→ 3')</u>	**Corresponding amino acid sequence**
120 (sense)	CGGT(C/G)AAGTC(C/G)AACATCGG	KSNIGHT
121 (anti-sense)	GC(A/G)ATCTC(A/G)CCCTGCGA(A/G)TG	HSQGEIA
122 (sense)	GT(C/G)GACAC(C/G)GC(C/G)TGCTC(C/G)	VDTACSS
123 (sense)	GG(C/G)AC(C/G)AACGC(C/G)CACGT(C/G)AT	GTNAHVI
124 (anti-sense)	CCCTG(C/G)CC(C/G)GGGAA(C/G)ACGAA	FVFPGQG

The selection of C or G where necessary in the third position of a codon reflects the very high overall G + C content of *S. cellulosum* (ca. 70%). Conditions for PCR were as follows: 0.01 M Tris-HCl (pH 9.0), 0.05 M KCl, 0.003 M $MgCl_2$, 0.1% Triton X-100, 200 μM of each primer, 2.5 U Taq DNA polymerase (Promega), 5.0% dimethyl sulfoxide (Sigma), and 0.01 μg of *S. cellulosum* chromosomal DNA in a 0.05 ml reaction volume. Reactions were carried out in a Perkin-Elmer Model 480 Thermocycler (Perkin-Elmer Corporation, Foster City, CA) under the following conditions: 94°C, 1 min; 50°C, 1 min, 72°C, 1.5 min for a total of 30 cycles. Each possible combination of sense and anti-sense primers were tried. A 650-bp and 350-bp fragment was amplified using oligos 120 + 124 and 123 +124, respectively. The sequence of the fragments were determined using the ALFexpress AutoRead kit to fluorescently label the DNA, which was analyzed on an ALFexpress sequencing apparatus (Pharmacia). The data indicated both PCR fragments possessed significant homology to polyketide synthase genes of Type I antibiotics. The 650-bp fragment was chosen for hybridization experiments.

The fragment was labeled with ^{32}P-dCTP using the NEBlot kit (New England Biolabs, Beverly, MA) and purified on a Bio-Spin 6 column (BioRad, Hercules, CA.). Duplicate blots were pre-hybridized in 3 X SSC (1 X SSC contains 0.15 M sodium chloride and 0.015 M sodium citrate, pH 7.0), 4 X Denhardt's solution (100 X is 2% Ficoll [Type 400], 2% polyvinylpyrrolidone, and 2% bovine serum albumin [Fraction V]), and 100 μg/ml sheared, denatured salmon sperm DNA; all reagents purchased from Sigma Chemicals, St. Louis. The labeled DNA was heated in a boiling water bath for 5 min to denature the strands, cooled on ice, and added to the pre-hybridization solution. The filters were incubated for at least 18 hr in a roller bottle hybridization oven. They were transferred to new bottle, then washed two times in 2 X SSC, 0.1% SDS at 70°C for 30 min (moderate stringency). The membranes were placed on Whatman 3MM paper to remove excess liquid, covered with Saran Wrap, and exposed to autoradiography film (Kodak X-OMAT LS) with two intensifying screens. The cassette was placed at -70°C and developed at appropriate intervals.

Approximately 100 colonies were seen to have hybridized on the duplicate filters. Fourteen of these were isolated from the master plate and grown in 4 ml LB + kanamycin medium for 20-24 hr, 37°C, 250 rpm. Plasmid DNA was prepared using the standard alkaline lysis method and digested with restriction endonuclease PstI. The digested DNA was electrophoresed on a 0.8% agarose gel in TAE for 3 hr at 100 V. Fragments were transferred to Duralon UV using the VacuGene XL vacuum blotting unit (Pharmacia) and the recommended alkaline denaturation protocol. Hybridization with radioactively-labeled PCR fragment and washing were carried out as described above. Two prominent types of cosmids were observed; one contained PstI fragments of ca. 7.0, 5.0, and 1.1 kb (pEPOcos6 and pEPOcos7) that hybridized to the probe; the other type had fragments of ca. 6.0 and 3.6 kb (pEPOcos8 and pEPOcos13) which were homologous to the probe. Restriction analysis confirmed that cosmids showing identical hybridization patterns had identical or overlapping inserts. PCR reactions using primers representing consensus sequences of Type I PKS genes were performed using the isolated cosmid DNA as template under conditions described above, except ca. 0.01 μg of cosmid DNA was included as template. Cosmids pEPOcos6 and pEPOcos8 amplified the 650-bp fragment seen when oligonucleotides 120 + 124 were used, while pEPOcos8 and pEPOcos13 supported amplification of an 1100-bp PCR fragment with oligos 122 and 124. The latter fragment was sequenced and confirmed to possess strong similarity to Type I PKS genes. These data confirm that the recombinant cosmids are related to each other and that all contain PKS-like genes.

2. Colony blot hybridizations of plasmid library in pZero2.1:

A 137-mm circle of Duralon UV membrane was placed on top of a 150-mm containing 75 ml LB agar + kanamycin. An aliquot of the plasmid library (representing ca. 2,000 recombinant colonies) in 0.5 ml LB medium was spread on the filter. The plate was incubated at 37°C for 20 hr. Colonies were replicated onto two fresh filters which were placed onto LB + kanamycin agar medium and incubated at 37°C for 6 hr. The filters were processed for hybridization as described in Section C.1. Out of 8 positive colonies detected, one contained a plasmid with a DNA region not encoded by either pEPOcos6 or pEPOcos8. This plasmid, called Sau4, was characterized in more detail.

3. Colony blot hybridizations of cosmid library in SuperCos:

The recombinant *E. coli* clones from the microtiter plates (see section 4. b) were used to produce two identical sets of hybridization filters in order to identify cosmids carrying PKS and PS genes. The recombinant clones were spotted onto 2 sets of 22 x 22 cm LB agar plates containing 50 µg/ml kanamycin. Each plate contained 384 clones therefore representing 4 microtiter plates. The clones were incubated at 30°C overnight. After pre-cooling for approximately 3 h at 4°C, 20 x 20 cm Hybond N^+ Nylon membranes (Amersham, Braunschweig, Germany) were placed onto the agar surfaces. After 2 min. the membranes were removed and placed for 15 min. on Whatman 3 MM paper (Whatman paper Ltd., Maidstone, England) soaked with denaturation solution (0.5 N NaOH, 1,5 M NaCl) before they were transfered onto Whatman 3 MM paper saturated with neutralization solution (1 M Tris-HCl, pH 7.5, 1.5 M NaCl). Subsequently the membranes were placed onto Whatman 3 MM paper soaked with 2 X SSC (0.3 M NaCl, 0.03 M sodium citrate, pH 7.2) for 10 min. The membranes were baked for 40 min at 85°C. Then, each membrane was overlayed with 5 ml Proteinase K solution (2 mg/ml Proteinase K in 2 x SSC) and incubated at 37 °C for 90 min. Finally, cell debris was removed by wiping the membranes with a Kimwipe pre-wetted with 2 X SSC.

As we were seeking in particular to identify biosynthetic pathways containing both PKS and PS genes, the following hybridization strategy was taken: The screening was initially focused on ketosynthase domains from type I PKSs and on the adenylation domain from PSs. Target-specific primers were used to amplify DNA fragments of the corresponding genes from chromosomal DNA of *S. cellulosum* by PCR. The fragments obtained were then cloned, sequenced and the deduced amino acid sequence compared to known ketosynthase and adenylation domains of PKS and PS, respectively. In a second step these PCR fragments were used as gene probes to detect recombinant cosmids of the *S. cellulosum* cosmid library.

Oligonucleotides based on conserved amino acid sequences of ketosynthase domains from various type I PKS were optimized for myxobacterial DNA by comparison to a known myxobacterial biosynthetic gene cluster (Schupp et al., J. Bacteriol. 177, 3673-3679 [1995]) resulting in primer
KS1Up (5´-
C/A)GIGA(A/G)GCI(A/C/T)(A/T)I(C/G)(C/A)IATGGA(C/T)CCICA(A/G)CAI(A/C)G-3´) and
KSD1 (5´-GG(A/G)TCICCIA(A/G)I(G/C)(T/A)IGTICCIGTICC(A/G)TG-3´).
PCR-primers TGD (5´-
T(A/T)(C/T)CGIACIGGIGA(C/T)(C/T)(G/T)IG(G/T)ICG-3´) and

LGG (5´-A(A/T)IGA(A/G)(G/T)(G/C)ICCICCI(A/G)(A/G)(G/C)I(A/C)(A/G)AA(A/G)AA-3´) directed to genes encoding adenylation modules have been described by Turgay et al. (Pept. Res. 7, 238-241 [1994]). PCR reaction mixtures with a final volume of 25 µl contained 0.1 µg template DNA, 0.2 U Taq DNA-polymerase (Gibco BRL, Eggenstein, Germany), 5 µmol dNTP, 5% dimethyl sulfoxide (Sigma), 1.5 mM $MgCl_2$, 25 pmol of each primer and the appropriate reaction buffer supplied by Gibco BRL. Chromosomal DNA of *S. cellulosum* was used as template. Additionally, chromosomal DNA of *Myxococcus fulvus* was used with PS primers. Reactions were carried out in an Eppendorf Mastercyler Gradient (Eppendorf, Germany) using the following conditions: denaturation 30 s at 97°C, annealing 30 s at 55°C, extension 60 s at 72°C for a total of 30 cycles. The formation of ca. 700 bp fragments using the KS primers and of ca. 350 bp fragments with the PS primers were confirmed by 0.8% agarose gel electrophoresis. Fragments of independent PCR reactions were ligated into vector pCR2.1TOPO using the TOPO TA Cloning kit (Invitrogen, Leek, The Netherlands) according to the manufacturer´s protocol and transformed into *E. coli* XL1-Blue. Sequencing of the resulting plasmids and analysis of the deduced amino acid sequence revealed three different KS fragments, designated pM008.4, pM008.6, pM008.7, one PS fragment (pAPs1) corresponding to *S. cellulosum* and one PS fragment (pDPs1) obtained with chromosomal DNA of *M. fulvus*. The PCR fragments were re-isolated by digestion with *Eco*RI from the plasmids pM008.4, pM008.6, and pM008.7, labeled, pooled and used as gene probes in hybridization experiments as described below. The same procedure was performed with the PS fragments of pAPs1 and pDPs1.

Hybridization with PKS and PS specific DNA probes (see above) was carried out using the DIG nonradioactive labeling and detection kit (Boehringer Mannheim, Germany) and performed according to the supplier´s manual using buffer containing 50% formamide. The membranes were hybridized in plastic bags containing approx. 10 ml of hybridization solution at 39°C overnight. Unspecific binding of probes was removed by 2 wash steps with 2 x SSC, 0.1% SDS at room temperature for 20 min. and one stringent wash step with 0.5 x SSC, 0.1% SDS at 60°C for 20 min. Detection of hybridizing DNA fragments was performed with the above mentioned system according to the manufacturer´s protocol using CSPD as chemiluminescent substrate. The signals were recorded by exposure of the treated membrane to Hyperfilm ECL (Amersham Life Science, Little Chalfont, England) which was developed in appropriate time intervals.

71 signals were detected with the PKS specific gene probe. On the duplicate filters 35 signals were obtained with the PS specific gene probe of which 7 were already known from the PKS hybridization experiment. These recombinant cosmids harbored PKS- and PS-encoding genes. In order to corroborate these results PCR experiments were performed with DNA of the 7 recombinant cosmids as template and PKS (KS1Up, KSD1) and PS specific primers (TGD, LGG) generating fragments of the expected size of approx. 700 bp and 350 bp, respectively (primers and reaction conditions see above).

A comparison of the restriction fragment patterns of the DNA from the 7 recombinant cosmids carrying PKS and PS genes digested by BamHI facilitated an arrangement of the cosmids in

3 groups. They were represented by cosmids designated A2 and A5. The remaining group was represented by pEPOcos6. Therefore, A2 and A5 represented good candidates for further DNA sequence analysis because they carry both PKS and PS genes.

D. Random "shotgun" sequencing of recombinant cosmids and plasmids

1. Library construction

a. pEPOcos6, pEPOcos8, A5, and Sau4: pEPOcos6 and pEPOcos7 were sequenced to completion, and contiguous sequence data and analysis for these overlapping cosmids is presented below for the "cos6 region" (cf. claims 7 and 9). Sequencing of cosmid A5, pEPOcos8 and plasmid Sau4 was taken to the point of large contiguous sequences (contigs) representing the *S. cellulosum* insert; sequence and analysis presented below (cf. claims 10 to 15).

Randomly sheared libraries were constructed for cosmids and plasmids of interest using a protocol similar to that of of Fleischmann et al., 1995 (Science <u>269</u>, 496) and modified in Fraser et al., 1995 (Science <u>370</u>, 397). Briefly, Qiagen-column purified cosmid DNA (~10 μg) was sheared to a size of approximately 2 kb and the DNA end-repaired using BAL31 nuclease. The DNA was gel-purified after electrophoresis through a 0.75% low-melting temperature agarose gel containing 0.5 μg/ml ethidium bromide in 1X TAE buffer run at 80 V for 2 hours. The volume of the low-melt agarose gel slice was estimated by adding the gel slice to a microfuge tube and weighing, then 0.1 vol. of 3 M sodium acetate (pH 7) was added and the agarose incubated at 60°C. The temperature was equilibrated to 37°C, and DNA extracted twice using an equal volume of buffered phenol (Life Technologies). The aqueous phase was transferred and extracted once with an equal volume of chloroform, then ethanol precipated by the addition of 2 vol. cold 100% ethanol. DNA was concentrated by spinning at 16,000 x g in a microcentrifuge. The DNA pellet was washed with 1 ml 70% ethanol and resuspended in 100 μl of 0.1X TE. The DNA was ligated to SmaI-digested, phosphatase-treated pUC18 vector (Pharmacia), and single insert recombinants isolated by gel-purification of the band containing vector plus a single insert, followed by T4 polymerase polishing, and a final intramolecular ligation of the vector-plus-single-insert DNA. This final ligation represents a library of highly random ca. 2 kb fragments that was used for shotgun sequencing of the ca. 40 kb cosmids or ca. 10 kb plasmids.

b. Cosmid A2: Cosmid DNA with inserts of *S. cellulosum* was isolated by an alkaline lysis procedure and purified with Macherey Nagel columns (Machery und Nagel GmbH und CoKG, Düren, Germany) using manufacturer's recommendation. Purified Cosmid DNA was sonicated, end-repaired using T4 DNA Polymerase (Boehringer Mannheim, Germany). After gel-purification fragments of a size of approximately 2 kb were ligated into SmaI-digested, phosphatase-treated pTZ18R vector (Pharmacia). The ligation represents a library of highly random ca. 2 kb fragments that was used for shotgun sequencing of the ca. 40 kb cosmid.

2. Sequencing and assembly

a. pEPOcos6, pEPOcos8, Sau4, and A5: DNA (1 μl of 100 μl total in the library) was transformed into *E. coli* by electroporation (20 μl of Electromax DH10B cells from Life Technologies) and cells spread onto LB plates containing 50 μg/ml ampicillin. After growth overnight at 37 °C, transformants (ca. 300-3000 CFU total) were tranfered to 96-well growth blocks and shaken overnight at 37°C in 1.3 ml LB medium with 50 μg/ml ampicillin. Templates were prepared from these cells by an alkaline lysis procedure (Qiagen QiaQuick Turbo Prep) to yield purified, double-stranded plasmid DNA. Cycle-sequencing of the plasmid templates was performed using universal forward and reverse primers and BigDye Terminator sequencing kits (Applied Biosystems), using the manufacturer's recommendations, then resolved using an ABI377 automated sequencer. Sequences were edited using Phred, then assembled into larger contiguous sequences using Phrap (Phil Green, University of Washington, St. Louis, MO).

b. **Cosmid A2:** DNA (1 µl of 20 µl total in the ligation) was transformed into *E. coli* DH10B by electroporation and cells were spread onto LB agar medium containing 50 mg/ml ampicillin. After growth for 18 hr at 37°C, transformants were transferred to 96-well growth blocks and shaken overnight at 37°C in 1.3 ml 2x YT medium with 50 mg/ml ampicillin. Templates were prepared from these cells by an alkaline lysis procedure (Qiagen Qiaquick Turbo Prep) to yield purified, double-stranded plasmid DNA. Cycle-sequencing of the plasmid templates was performed using universal forward and reverse primers and Big Dye Terminator sequencing kits (PEBiosystems) or Thermo Sequenase fluorescent labelled primer cycle sequencing kit (Amersham Pharmacia Biotech) using the manufacturer's protocols. In the shotgun phase of a cosmid, identical amounts of samples were sequenced either by dye-primer or dye-terminator chemistries (Pharmacia, PE Biosystems). Data were collected using Licor and ABI 377 automated sequencers and assembled with the GAP4 program (Bonfield, Smith, Staden, Nucl. Acids Res. 23, 4992-4999 [1995]). Gaps were closed using custom made primers (MWG-Biotech) on plasmid templates or PCR products in combination with dye-terminators.

E. Bioinformatic Methods

1. Open reading frame (ORF) identification

ORFs were identified in the pEPOcos6 region using the OMIGA 1.1.2 (GCG 0.4D) program from Oxford Molecular Limited. Default values were used (Standard genetic code, all ORFs over 50 bases) to generate ORFs; analysis of these results lead to the list of 14 highest quality ORFs as defined in claim 9. Other ORFs, genes, or genetic elements may be found in the pEPOcos6 insert that have not yet been annotated. In addition to hand-editing of the OMIGA-generated data, the MAGPIE automated genome analysis tool:

(http://genomes.rockefeller.edu/magpie/magpie.html)

was used to identify genes for all the sequenced cosmids and plasmids. ORFs identified in this manner are presented as both nucleotide and peptide files below.

For cosmids A2 and A5, ORFs have been identified within the DNA sequences of A5 (contigs 10, 11, 12) and of A2 using the FramePlot analysis program from Ishikawa and Hotta (FEMS Microbiol. Lett., 174, 251-253 [1999] public available under [http://www.nih.go.jp/~jun/cgi-bin/frameplot.pl] which is based on positional base preference in codons typical for organisms having genomes with a high G + C content (Bibb et al., Gene 30, 157-166 [1984]). Default parameters using ATG and GTG as start codons were used. The deduced amino acid sequence of predicted ORFs were compared with protein databases (GenBank, CDS translations, PDB, SwissProt, PIR, PRF) using BLASTP (Altschul et al., Nucleic Acids Res., 25, 3389-3402 [1997]). Additionally, high scoring amino acid sequences were analyzed using the Pfam program [http://www.sanger.ac.uk/Software/Pfam/], which identified specific domain structures of the submitted proteins (Bateman et al. Nucleic Acids Res., 27, 260-262 [1999]).

2. BLAST searches

BLASTP2 similarity searches were performed using the peptide files from the above ORF identification strategy as query sequences. Searches were performed using the in-house Bioinformatics BLASTP2 (Version: BLASTP 2.0a19MP-WashU) web page at the Bristol-Myers Squibb Pharmaceutical Research Institute (allows BlastN2, BlastP2, BlastX2, TblastN, and TBlastX searches). In addition, peptide files generated by the MAGPIE analysis were automatically searched using a FASTA algorithm.

3. Best match and probable identification

Analysis of the BLASTP2 and FASTA output led to an assignment of a best match and probable function. The best match was usually the top scoring match, although sometimes another match was given because it was a more relevant homolog, or no match was found with a significance greater than >e-4. Probable function represents the best estimate of function given the initial analysis of the BLAST data and the published literature regarding the best match, and may not necessarily represent the true function of the gene product (hypothetical proteins are of unknown function). A higher probability score indicates a higher liklihood that the probable function corresponds to that of the best match; e.g., the polyketide synthase matches are all above e-100, and given the very high significance scores are presumed to function as polyketide synthases (as are the high scoring peptide synthetases).

The following is a summary of the sequence data from the pEPOcos6 region, pEPOcos8, A5, Sau4 and A2.

a. Data from pEPOcos6 region:

Summary: A large PKS/PS cluster spanning multiple cosmids. An IS element (designated IS-Sc1 here) is found in the cluster - this may be a potential tool for genetic analysis of *Sorangium*.

Statistics: Sequence was assembled from over 2000 random sequences (forward and reverse reads of the ca. 2 kb cloned fragments derived).

47,713 nucleotides of contiguous sequence (no pFD666 vector included)

DNA sequence data are as defined in claim 7.

Note: pEPOcos6_ORF7 sequences (cf. claim 9): the predicted N-terminus of ORF7 shows 145 nucleotide overlap with ORF6.

Note: pEPOcos6_ORF8 sequences (cf. claim 9): >pEPOcos6_ORF8.seq ("ORF9_up" in Fig.2)

67.3% G+C

Table 3 shows ORF data summary. Note: pEPOcos6_ORF1.seq is truncated at its 5' end; correspondingly pEPOcos6_ORF1.pep is truncated at its N-terminus.

b. Data from pEPOcos8 region:

Summary: Two PKS genes found on a cosmid. A Tn1000 insertion is also found (occurred during *E. coli* propagation). No peptide synthetase genes were found; one P450 hydroxylase was identified.

Statistics: 1952 random sequence reads from the pEPOcos8 library were assembled using phrap, with 1024 of the sequences assembling into 57 contigs. 12 of these contigs were chosen (totaling 56,537 bp) which each contained >6 reads and consisted of about 1000 bp or more. The sequences of these 12 contigs and the associated ORFs are given below.

DNA sequence data from contigs are as definded in claim 10. Table 4 shows more data.

pEPOcos8 protein data are as defined in claim 11, i.e. for selected ORFs (polyketide synthase, peptide synthetases, or ORFs with high similarity to known genes).

c. Data from cosmid A5 insert:

Summary: A cluster of PKS and PS genes found on the cosmid. Other genes possibly involved in this secondary metabolite production include a downstream lipoxygenase gene higly similar to eukaryotic orthologs.

Statistics: 880 random sequence reads from the A5 library were assembled using phrap, with 530 of the sequences assembling into 12 contigs. 3 of these contigs were chosen (totaling 41,556 bp) which each contained >100 reads and consisted of about 9000 bp or more. The sequences of these 3 contigs and the associated ORFs are given below.

DNA sequence data from contigs are as defined in claim 12. Table 5 shows more data.

Protein sequence data from selected A5 ORFs are as defined in claim 13.

d. Data from plasmid Sau4 insert:

Summary: Insert contains PKS genes on two large contigs - most similar to the soraphen PKS gene from Sorangium.

Statistics: 565 random sequence reads from the Sau4 library were assembled using phrap, with 84 of the sequences assembling into 18 contigs. 2 of these contigs were chosen (totaling 6596 bp) which each contained >10 reads and consisted of about 1000 bp or more. The sequences of these 2 contigs and the associated ORFs are given below.

DNA sequence data from plasmid Sau4 contigs are as defined in claim 14. Table 6 shows more data.

Protein sequence data from selected plasmid Sau4 ORFs are as defined in claim 15.

e. Data from cosmid A2

Table 7 shows ORF data summary

F. Construction of suitable recombinant expression vectors

1. Expression in Myxobacteria

Heterologous expression of the ORFs shown in Figure 1 is performed by using a derivative of plasmid pSUP102 (Simon, R., Priefer, U., Pühler, A., Methods in Enzymology (1986), vol. 118, pp. 643-659). In this plasmid the gene for chloramphenicol resistance is changed for a cassette comprising the gene for streptomycin resistance and the promoter element of the Tn5 transposon. Short homologous genomic DNA segments from the host organism are ligated with the DNA sequences of Figure 1 and with efficient regulatory elements into, for example, the EcoRI restriction site of the vector. Following amplifiction of the vectors in *Escherichia coli* the DNA is transfered by electroporation of the host cells or by conjugation with *Escherichia coli* S17-I (Simon, R., Priefer, U., Pühler, A., Biotechnology (1983), vol. 1, pp. 784-791).

By means of the tetracycline or streptomycin resistance, respectively, mediated by the vector the host cells are checked for integration of recombinant plasmid DNA into the chromosome by homologous recombination.

2. Expression in Streptomyces cells

Heterologous expression of the ORFs shown in Figure 1 is performed by using bifunctional Strepomyces-Escherichia coli cosmids pKU206 and pOJ466.

3. Expression in Escherichia coli cells

Heterologous expression of the ORFs shown in Figure 1 is performed by using "bacterial artificial chromosomes", cosmids (for example Supercos, Stratagene GmbH, Heidelberg) and T7 expression systems (Stratagene GmbH, Heidelberg; New England Biolabs Schwalbach, FRG). Expression of recombinant enzymes occurs in Escherichia coli cells constitutively expressing phosphopantetheinyl transferase required for the formation of holoenzyme polyketide synthetases and polypeptide synthetases.

Table 3. pEPOcos6 region gene annotation summary (continued).

GENE	5'	3'	# BP	#AA	MW (DA)	BEST MATCH (SCORE)	PROBABLE IDENTIFICATION
ORF1	1 [a]	731	732 [b]	244	26218	ACDA_BACSU (e-22)	acyl-CoA dehydrogenase
ORF2	731	2275	1545	515	56317	AF047828 (e-94)	peptide synthetase
ORF3	2301	3356	1058	352	39092	U57065 ZmaR (e-29)	antibiotic resistance gene
ORF4	3412	4050	639	213	24819	(no match)	hypothetical protein
ORF5	4248	12101	7854	2618	283282	PKSK_BACSU (e-190)	polyketide synthase
ORF6	12097	14322	2226	742	80545	Z83857 ppsE (e-111)	polyketide synthase
ORF7	14178 [c]	16451	2274	758	82007	(no match)	hypothetical protein
ORF7.1 [d]	15866	14196	1671	557	61320	Y4HP_RHISN (e-28)	hypothetical protein
ORF7.2 [d]	16507	16154	354	118	13102	Y4HO_RHISN (e-13)	hypothetical protein
ORF7.3 [d]	16815	16507	309	103	11079	Y4RG_RHISN (e-13)	hypothetical protein
ORF8	16830	17630	801	123	29633	PKSL_BACSU (e-2)	polyketide synthase
ORF9	17628	25244	7617	2539	279286	pir|S73015 (e-160)	polyketide synthase
ORF10	25235	27877	2643	881	97101	AF047828 (e-111)	peptide synthetase
ORF11	27867	32498	4632	1544	165084	AF091251 (e-167)	polyketide synthase
ORF12	32498	39922	7425	2475	267116	AF040570 (e-223)	polyketide synthase
ORF13	40031	45559	5529	1843	71258	AF091251 (e-119)	polyketide synthase
ORF13.1	45599	46018	420	140	14590	(no match)	hypothetical protein
ORF14	46015	47577	1563	521	55671	PKN1_MYXXA (e-34)	Serine/Threonine-Protein kinase

[a] The predicted ORF1 gene and gene product is truncated due to cloning of the DNA into the cosmid vector.

[b] 731 is the last nucleotide of the last amino-acid-encoding codon; 732-735

is TGA (termination codon). Termination codons have been excluded in the present annotation.

[c] note: the predicted N-terminus of ORF7 shows 145 nucleotide overlap with ORF6

[d] note: ORFs on opposite strand - make up a transposable element from bp 16863-14130 (2733 bp, 11 bp terminal inverted repeat) that is similar to IS1131 from *Agrobacterium tumefaciens* (IS-66 like element, 2773 bp, 4 ORFs, 11 bp inverted repeat)

Table 4. pEPOcos8 assembly analysis summary (continued).

a. pEPOcos8 assemblies

ASSEMBLY	SIZE(BP)
Contig43	1017
Contig44	1246
Contig48	978
Contig49	1969
Contig50	2877
Contig51	2319
Contig52	1883
Contig53	4871
Contig54	7257
Contig55	5021
Contig56	10945
Contig57	16154

b. selected ORFs from Contig 56 and 57 of pEPOcos8

gene	5'	3'	# bp	#aa	best match [score]	probable identification
Contig56_003	3	8675	8673	2890	ERY2_SACER (e-300+)	polyketide synthase
Contig56_027	10784	8682	2103	700	pir|I60218 (e-300+)	transposon Tn1000 (*E. coli*)
Contig57_001	92	1210	1119	372		
Contig57_002	222	7001	6780	2259	AL021899 pks12 (e-267)	polyketide synthase (M. tuberculosis)

Contig57_026	7424	6975	450	149		
Contig57_027	7116	8375	1260	419	CPXK_SACER [e-50]	Cytochrome P450 monooxygenase
Contig57_030	9074	8337	738	245		
Contig57_031	8412	9080	669	222		
Contig57_032	8432	9667	1236	411		
Contig57_033	9724	8600	1125	374		
Contig57_037	10621	9755	867	288		
Contig57_039	11085	10618	468	155		
Contig57_040	10687	11091	405	134		
Contig57_041	11887	10904	984	327		
Contig57_043	11520	12998	1479	492	AF072709 [e-42]	unknown ORF (S. lividans)
Contig57_044	13730	12018	1713	570		
Contig57_047	13093	13797	705	234		
Contig57_048	14064	13138	927	308		
Contig57_049	14371	13535	837	278		
Contig57_051	13900	14394	495	164		
Contig57_052	14036	14440	405	134		
Contig57_053	14678	15715	1038	345		
Contig57_056	15211	15900	690	229		

Table: 5. A5 assembly analysis summary (continued)
a. pEPOcos8 assemblies

contig	bp	ORF	5'	3'	bp	G+C content pos.3	best match [score]	probable function	identified domains using Pfam
10	9435	1	2861	582	2280	88.6 %	AAC44128 (e-170]	peptide synthelase [Saframycin]	AMP-binding aa 68-512; pp-binding aa 631-697
		2	3525	2902	624	87.0 %	no match		
		3	6420	4195	2226	95.1 %	P38370 [3e-04]	OmpA like	
		4	7734	7135	600	98.0 %	P39760 [2e-21]	unknown	
		5	9434	7950	>1495	96.2 %	no match		
11	18023	1	2526	751	1776	97.1 %	CAB38518 [e-134]	peptide synthetase	AMP-binding aa 118-520;
		2	3586	2501	1086	94.5 %	CAB38516 [e-12]	peptide synthetase	DUF4 (domain of unknown function, found to the carboxy side of pp-pinding sites) aa 5-360;
		3	5210	3579	1632	97.8 %	S53999 [e-102]	peptide synthetase [Pyoverdine]	pp-binding aa 1-30; DUF4 aa 51-442; conflict. ca. 20 aa at N-terminus are missing (conserved serine residue)
		4!	6261	3811	2649	64.4 %!!	CAA11039 [7e-55]	polyketide synthase [Rifamycin]	conflict!!
		5	8087	6258	1830	94.8 %	CAB06094 [e-144]	polyketide synthase (Phenolpthiocerol Mycob.)	ketoacyl-synthase aa 14-439;
		6	17320	8090	9231	92.8 %	AAD04757 [e-180]	peptide synthetase [Lichenysin]	AMP-binding aa 371-758; pp-binding aa 840-904; DUF4 aa 937-1319; AMP-binding aa 1438-1825; pp-binding aa 1907-1071; DUF4 aa 1999-2388; AMP-binding aa 2485-2889; pp-binding aa 2972-3004; pp-binding aa 3027-3046;
12	15898	1	1	1545	>1545	97.5 %	AAC68816 [4e-74]	peptide synthetase (FK506]	DUF4 aa 45-426;
		2	2883	1549	1335	94.4 %	P43492 [e-07]	cytochrome P450 enzyme	
		3	4659	2911	1749	92.5 %	NP 001130 [5e-56]	lipoxygenase	
		4	5896	7086	1191	95.2 %	no match		
		5	7094	7822	729	90.1 %	no match		
		6	8974	7943	1032	87.8 %	no match		
		7	12001	11252	750	92.8 %	no match		
		8	13533	12181	1353	93.3 %	no match		
		9	15897	13981	>1917	93.9 %	CAA19149 [e-09]	regulator	
	43356								

Table 5.

b. selected ORFs from cosmid A5

gene	5'	3'	# bp	#aa	best match [score]	probable identification
Contig10_001	2861	582	2280	759	U24657 saframycin (e-155)	polyketide synthase
Contig10_006	2880	3323	444	147		
Contig10_007	3876	2902	975	324		
Contig10_008	2953	3957	1005	334		
Contig10_009	4357	3299	1059	352		
Contig10_013	4180	6585	2406	801		
Contig10_016	7033	5600	1434	477		
Contig10_017	7043	6138	906	301		
Contig10_018	6500	7219	720	239		
Contig10_019	6954	7295	342	113		
Contig10_020	6982	7797	816	271		
Contig10_021	7819	7040	780	259		
Contig10_023	7415	8029	615	204		
Contig10_024	7794	9435	1642	548		
Contig10_027	9435	8806	630	209		
Contig11_001	770	417	354	117		
Contig11_002	2526	751	1776	591	AL035640 (e-113)	peptide synthetase
Contig11_004	1033	2787	1755	584		
Contig11_005	3500	1977	1524	507		
Contig11_007	3586	2501	1086	361		
Contig11_008	2507	3814	1308	435		
Contig11_011	5213	3579	1635	544	AF047828 syringomycin (e-86)	peptide synthetase

Contig11_012	6459	3811	2649	882	AJ223012 (e-52) rifamycin	polyketide synthase
Contig11_016	6511	5210	1302	433		
Contig11_018	5551	6735	1185	394		
Contig11_021	9899	6258	3642	1213	Z83857 ppsE M. tuberc (e-129)	polyketide synthase
Contig11_026	17329	8090	9240	3079	AF047828 syringomycin (e-300+)	peptide synthetase
Contig11_048	16733	17365	633	210		
Contig11_049	17397	17723	327	108		
Contig12_001	1	1545	1545	514	AF082100 (e-63) (Streptomyces)	peptide synthetase
Contig12_002	1368	1	1368	456		
Contig12_003	3	1655	1653	550		
Contig12_005	2317	1361	957	318		
Contig12_006	2883	1549	1335	444		
Contig12_007	1777	4926	3150	1049		
Contig12_009	4659	2911	1749	582	LOX5_MOUSE (e-54)	ARACHIDONATE 5-LIPOXYGENASE
Contig12_011	4523	5065	543	180		
Contig12_012	4638	5231	594	197		
Contig12_013	4942	5520	579	192		
Contig12_014	6056	5541	516	171		
Contig12_015	5765	6373	609	202		
Contig12_016	5896	7086	1191	396		
Contig12_017	7095	5899	1197	398		
Contig12_018	5955	7331	1377	458		
Contig12_020	7549	7010	540	179		
Contig12_021	7094	7822	729	242		
Contig12_022	7995	7318	678	225		
Contig12_023	8408	7716	693	230		
Contig12_024	7916	9550	1635	544		

Contig12_029	10751	9732	1020	339
Contig12_030	9795	11132	1338	445
Contig12_033	10880	12124	1245	414
Contig12_035	12184	11252	933	310
Contig12_036	12079	14098	2020	674
Contig12_039	14098	12248	1851	616

Table 6. Sau4 assembly analysis summary.

a. plasmid Sau4 assemblies

Assembly	size(bp)
Contig17	2581
Contig18	4015

b. selected ORFs from cosmid A5

gene	5'	3'	# bp	# aa	best match [score]	probable identification
Contig17_001	2485	1	2485	829	U24241 *Sorangium* [e-213]	polyketide synthase
Contig18_002	2	1510	1509	502	U24241 *Sorangium* [e-105]	polyketide synthase
Contig18_010	1494	4015	2522	841	U24241 *Sorangium* [e-245]	polyketide synthase

Table: 7. ORF data summary from A2 insert

ORF	5´	3´	bp	G +C content pos.3	best match [score]	probable function	identified domains using Pfam
1	1666	1	>1666	94.4 %	P54744 [5e-37]	regulation [serine/threonine protein kinase]	pkinase aa 47-294:
2	1605	3338	1734	90.7 %	no match		
3	6100	3398	2703	98.0 %	CAA15124 [>e-180]	acyl-tRNA synthetase	tRNA-synthase aa 27-694;
4	7111	6374	738	94.7 %	AAC32457 [2e-17]	monooxygenase	oxidoreductase FAD/NAD-binding domain aa 110-227;
5	9590	8433	1158	74.4 %	CAB42045 [9e-67]	aminotransferase	aminotran_1 aa 2-385;
6	11393	9855	1539	85.5 %	AAD21754 [e-113]	L-dopa decarboxylase	pyridoxal deC aa 46-411
7	13656	12712	945	86%	CAB41201[7e-41)	oxidoreductase	short chain dehydrogenase domain aa 53-240;
8	15374	19984	4611	87.6 %	CAB06094 [<e-180]	polyketide synthase	ketoacyl-synthase aa 13-438; acyl transfase aa 533-854; short chain dehydrogenase domain aa 1159-1357; pp binding aa 1451-1515;
9	20003	27889	7887	88.0 %	AAC80285 [<e-180]	polypeptide synthetase	DUF4 aa 17-409; AMP-binding aa 507-905; pp-binding aa 991-1054; DUF4 aa 1067-1466; DUF4 aa 1544-1944; AMP-binding aa 2041-2439; pp-binding aa 2525-2589;
10	28251	29402	1020	86.6 %	BAA13079 [2e-44]	peptidase	
11	31720	30401	1320	98.9 %	no match		
12	31982	32932	951	95.3 %	CAB09733 [2e-63]	sigma factor	Sigma70 ECF aa 17-83;
13	33128	33613	486	95.7 %	no match		
14	33661	34077	417	94.2 %	CAA19900 [e-37]	regulation	
15	35611	35255	357	80.8 %	CAA19859 [3e-15]	regulation	HTH aa 5 11-109;
16	37856	35730	>2127	96.8 %	BAA17885 [2e-50]	regluation	response reg aa 46-159; signal aa 326-542; response reg aa 590-703;

Claims

Claims

1. DNA sequence, the expression products of which cause an enzymatic biosynthesis, a mutasynthesis or a partial synthesis of polyketide or heteropolyketide compounds or are involved therein.

2. DNA sequence according to claim 1, wherein the polyketide or heteropolyketide compounds are epothilones.

3. DNA sequence according to any of the preceding claims, wherein the DNA is derived from myxobacteria.

4. DNA sequence according to any of the preceding claims, wherein the DNA is derived from Sorangium strains.

5. DNA sequence according to any of the preceding claims, wherein the DNA is derived from Sorangium cellulosum.

6. DNA sequence according to any of the preceding claims, wherein the DNA is selected from the group consisting of:

(a) the following DNA sequence:

Seq ID No 1 (A2 cosmid)

GGATCGCGGCGCCCTCGCGCTGCTCCTCGAGCGTGCGGAGGAACTCCCACGCCAGGCGCGACT
TGCCGAGGCCAGGCGCGCCCACCACCACCACCGCGTTCGCGGAGGGCTCGTCGACGCAATGGC
GCCACTCGGTCGCGAGCTGCGAGAGCTCGCGCTCCCGCCCCACGCAGGGCGTCGGCTTGCCGA
GGAGCCGTGGGACGGCATCCGGCTCCTCCTTCGGGCCGCGAAGCCAGCACCCTCCGGGCCCCT

GTACCGTCTCGAAGCGGCTCGCGAGCAGGCTGGCCGTCGCGTCGTCGAGCCGGATCTCCGGCG
GCGACAGGCCATCTCGCCCGGCGATGAGCTGCGCGACCCGATCGACCAGCTCGCCGACCGGCA
GCCTCGCCTCGACCTCGGCCAGCCCTGTCGCGACGGACACGGGCACGCCTCCGAGCGCCGCCC
GCAGCGCGAGGGCGCAGTGGGCCGCCCGTGTGGCGAGATCCGTGGGCGACTCGGCGCCGGACA
GCGCGACGAGCCACCAGCGCGCTTGCAGCCGATCGAGGCGCCCGCCGTGGCGCGCCGCGATGT
CCCGCAGCGCCTCGGCCCGCGCGGCGCCGTCGTCCTCCGAGAGCGTGGCGCCGGCCTCGGCGC
CGCCGTCTTCGGCCAGGATGACGCACATCACCTTGCGCTCGGCCGTCGTGATCGCCTCGCCCG
GCGCGGCCGGCGCCGCGACCGCGCTCGCCCCGATCGAGAGCCCCTCGCCGGCCACGGCGGCGA
GCTCCGCCGCGGCGGCGGCGCCGTCGCGCGGCCGCTCTCCCGCGTTCTTCGCCAGCATCCGCG
CCACCAGGCGCTCGAGCGGCTCCGGGATACCGTCGCGGAGCTCCCCGAGCCGCGGCGGCTCTT
CCAGGACGACCCGCATCAGGAGCGCGAGCGCGCTGTTGCCGAGGAACGGCGGGCGCCCCGCGA
GGCACTGGAACAGCACGCACCCGAGCGCGAACACGTCGGCCCGGGCGTCGACCGGCGCGTCGC
CGCGCACCTGCTCGGGCGCTATGTACCCGGGCGTGCCGAGCACGGCCCCGGGCGACGTGAGGG
TCGGCGCGAGCCGGAGGTGGCGCGCGATGCCGAAGTCGAGCAGCGTGACGCGCTCGACCGCGC
CGCCCACGAGCATCAGGTTGCTCGGCTTGAGGTCGCGGTGAACGACGCCGAGCCAGTGGATCG
CGCCGAGCGTCGTGGCCACGCGCGCGGCCAGCGCCACGCTCTCGGCCAGCGTGAGCGGCGCCC
CGGCGAGCCGCTCCTCCAGGGTCACGCCGTCGAGCCACTCCATGGCCAGGTACGGCCGCCCTG
CGCCGGTCACCCCGTGCGCCACGTACTGCACCACGCCGGGCAGCCGGAGCGTCACGAGCGCCT
CCGCCTCCCGCGCGAACCGGCGCAGGTCGTTGGCGCTCGCGCCCTGCAAGACCTTGAGCGCGA
CCGCCTGCCCGGACACCCGGTCGCGCGCCCGGTACACGTCCCCCATCCCGCCGGAGACGGCGA
GCCGCTCGATCTCGAAACGATCCTCGATCACATCCGCTGCGCGCATGGCGGTGCCAATGTACT
CCGCGCGAGCCTCGGGCCCCCGCGCGTAAGTGCGGCCCTGCGCCCGGTTGAACGCCAGCCCGA
GCGTGACCGCCTCGCGCTCGGGATCCACGGCCGCCGGATCGGTCCACGCCTCGACGAGCGCCT
GCGTTGAACAACCCGCCACCGGGCGCACGCAGCCGGCATCGCCGCGCTGGCCACCCGGCGCTG
CCGCCCTTAGGCTCACCTCCGCGATGCCCCGCTGGTTCAACACGGCAGGTCCCTGCAACCCGG
CCGATCACTACATGCTCCCGGCCGAGGAGCGCTTGCCCGCAGTGCGCGATCTGGTCGATCGCA
AGGCCTACTTCGTCCTGCACGCCCCGCGGCAGATCGGCAAGACGACCTCGCTGCGCACGCTCG
CCCAGGATCTCACGGCCGAAGGGCGCTACGTGGCCGTCCTCGTCTCGGCGGAGGTCGGCGCCC
CCTTCTCTGACGATCCCGGCGCGGCCGAGCTCGCGATGCTCGCAGAATGGCGCGGCACCGCCG
GCGCGCAGCTCCCCGCCGATCTGCGGCCGCCACCGTTCCCCGATGCGCCCGCCGGTCAGCGCA

TCGGGGCCGCCCTGCGCGCCTGGGCTCAGGCCGCGCCGCGCCCGCTCGTCGTCTTCCTCGACG
AGGCCGACGCCCTGCGCGACGCGACGCTCGTCTCCCTATTGCGCCAGATCCGCAGCGGCTATC
CCGACCGCCCGCGTGACTTCCCGCACGCGCTCGCCCTCGTCGGCCTGCGCGACGTGCGCGACT
ACAAGGTCGCGTCGGTCGACAGCGGCAGGCTCGGGACGTCGAGCCCCTTCAACATCAAGGTCG
AGTCGCTCACGCTGCGCAACTTCACCCGCGACGAGGTCGCAACACTCTACGCTCAGCACACGG
CCGAGACCGGTCAGGTCTTCCGGCCGGACGCCGTGGACCGCGCCTTCGAGCTCACCCAGGGCC
AGCCGTGGCTCGCCAACGCGCTCGCCCGCCAGCTCGTCGAGGTCCTCGTCAAGGACCGCGCGC
AACCCATCACGTCTGCGAACGTCGATCGCGCCAAGGAAATCCTCATCGAGCGGCAGGACACAC
ACCTCGACAGCCTGGTGGATCGGCTGCGCGAGCCGCGCATCCGCGCGGTGATCGAGCCGATGC
TCGCCGGCACCGCGTTGCCGAGCGTGCCCCCCGACGACCTTCGTTTCGCGATCGACCTCGGCC
TCGTGCGCATGACCGCGGAGGGCGGCCTCGACGTCGCCAACCCCATCTACCGCGAGATCATCG
TCCGCGAGCTCGCGTTCCCGATCCGCGCCTCACTCCCCCAGATCAAGGCCACGTGGCTCACGC
AGGACGGCCGCCTCGACGCGGACCGCCTGCTCGACGCCTTCCTCTCCTTCTGGCGCCAGCACG
GCGAGCCGCTCCTCGGCGCCGCGCCCTACCATGAGATCGCCCCGCACCTCGTGGTGATGGCCT
TCCTCCACCGCGTGGTGAACGGCGGTGGCACCGTCGAGCGCGAGTACGCCATCGGCCGGGGCA
GGATGGATCTCTGCGTTCGTTACGCGGGCGAGACGCTCGCGATCGAGCTCAAGGTCTGGCGAG
ACGGCCGCCCCGATCCCGTCGCCGAGGGGCTCGCCCAGCTCGACGAGTACCTGGCCGGCCTGG
GCCTCGATCGCGGATGGCTCATCCTCTTCGACCAGCGCTCCGGACAGCCCCCCATCGCCGAGC
GCACGCGCCGCGAGCGCGCGCTCTCCCCCGCCGGCCGCGAGGTCGCCGTCATTCGCGCCTGAG
GGAGCTCGCCGCGCGGCGAGCGCCCTCCACGAGGGCCGGGCCACCTCGGACAGCGTCTCTACT
CCTCCGAGGCCGCCGCGCCCCCCGCCCCGGCCGCCGCCGCCGCCGGCTCCAGCTCGCAGC
GCACCACCAGGACCTCGCCATCCGCGAGCTCCGGCCGCTCCACGAGCGCGTGCGCGCCCGCGC
GCACCGCCGTGAGCACGTCTCCCAGCGCCGGCTTCAGCCGCGCCAGCGTCGCGGCGTTCGCCC
CGAGCGCGAGGTCGGTCACGACGCGCCCCACGCTCGCGCCGAGCTCGCTCTTGCGCTTGTTGA
CCGCCGCCATCGCCGCCGCCGCCAGATCCAGGAGCCCCGGATCCGAAGGCGCCGCGACCGCCG
CGAAATCCGCCGCTGAAGGCCACTTCGCCCGGTGGATCGAGGTATCGCCCGTCTCCTCCGCGT
ACACCCAGCGCCAGACCTCGTCGGTGATGTACGGCAGGACCGGCGCGAACAGCCGCAGCAGCA
CCGACAGCCCGAGCCGCAGCGCCGCCACCGCCGAGCCGCGCGCCGCCTCCCCGGCGCCGCCCT
CGCCGCGCGCCCGCGCCTTCGCGAGCTCCAGGTAGGCGTCGGTGAACCAGCGCCAGAAGAAAT
CCTCGGTCCGCTCGAGCGCCGCCGCCGAACTCGTGCTCGTCGAACGAGCGCGTCGCGTCGTCCA

```
CCACGGCCGACAGCTTGTGCAAGAGCGCCCGGTCGAGCTCCTCGGAGATCGGGTGGACCTCCG
CCGACTGGCTGAGCACGTACTTGCTCGCGTTCCAGATCTTCGTGACGAGCCGCTTGCCGATCT
TCAGCACCTTCTCGTCGAACGCCGTGTCCGTGCCGAGCCGCGCGCTCGCCGACCAGTAGCGGA
CCGCGTCCGAAGAATACGTGTCGAGCAGGTGCATCGGCGTGACGACGTTGCCCTTGCTCTTCG
ACATCTTCTTGCGATCCGGATCGAGGATCCACCCGGAGATCGCGACGTGGTGCCACGGGACCG
ACGACTCGTGCAGCATCGCCTTCGCGATCGTGTAGAACGCCCACGTCCTGATGATGTCGTGGG
CCTGCGGCCGCAGATCGGCCGGGAAGAGCCGCGCGTGGCGCGCCGGATCGTCCCCCCAGTGAG
AGCTGATCTGCGGCGTGAGCGAGCTCGTGAACCACGTGTCGAAGACGTCGGACTCGGCGGTGA
AGCCGCCGGGCTGGTCCCGCTGCGACGCCTCGTACCCGGGCGGCACGTCGACCGTCGGGTCGA
CCGGGAGCATCTCGCGCGTCGCGAGCAGCGGCCGGCTGTGATCCGGGTTGCCCTCGGCGTCGA
GCGGATACCAGACCGGGAACTGCACGCCGAAATACCGCTGGCGGCTGATGCACCAGTCACCCT
GGAGCCCCTCGGTCCAGTTGCGGTACCGGAGGCGCATGAAATCCGGGTGCCACTTGATCTTGT
CGCCGTATTCGAGGAGCTCGGCCTTCTTGTCGGCGAGCCGGACGAACCACTGCCGCGTGGGCA
CGAACTCGAGCGGCTGGTCGCCCCGCTCGTAGAACTTCACCGCGCGCTCGATCGGCCTCGGCT
CGCCCCGCAGCGCCGGCCCCCGGCCGGGCGCCGCCGCGTGCTCCTCGCGGCGGAGCAGCTCGA
CCACCGCCGCGCGCGCCTGCTTCACCCCCCTGCCCTGGAGCGGCGCATACGCGGCGTTGGCCG
CGGCCGGGTCGCGGCTCTCCCACGCGCCCTCGCCGAACGTCACCGGCAGGACACGGCCGTTCT
TGCCGAGCATCTGCCGGAGCGGAGCTTCTGCTCCCGCCACCAGATCACGTCGGTCGCGTCGC
CGAAGGTACAGACCATCAGGATGCCCGTGCCCTTCTCGCGATCCACGAGCGGGCTCGGGAAGA
TCGGCACCGGCGCGCGGAAGATCGGGGTGAGCGCCGTCTTGCCGAAGAGGTGCTGATACCGCG
GGTCCTCCGGGTGCGCCGTGACGCCGACGCAGGCCGCGAGCAGCTCCGGGCGCGTCGTGGCGA
TGACGAGCTCCTCGGCCGTCCCCTCCACCGCGAACGCGATGTCGTGGAACGCGCCCGATTGCG
GGCGATCCTCGACCTCCGCCTGGGCGACCGCGGTCTGGAAATCGACGTCCCACATCGTCGGCG
CGAAGACCGAGTAGAGGTGGCCCTTCTCGTGGAGATCCAGGAACGACAGCTGCGCCGTCCTGC
GGCAGTGATCATCGATGGTGGCGTACTCGTTCCGCCAGTCGACCGAGAGGCCCACCCGGCGGA
AGAGCGCCTTGAAGACCTGCTCGTCCTCGCGCGTGACCTTGTGGCAGAGCTCGATGAAGTTGG
GCCGCGACACGATGCGCGGCGGCTCCTTCTTGATCGTCTCCGGCGCGGCCTGCGGCAAGGTCA
GGCCGCGCTCGTACGGCGTGCGCACGTCGGTGCGGACGTGGAAGTAGTTCTGCACGCGCCGCT
CGGTGGGCAGGCCGTTGTCGTCCCAGCCCATCGGGTAGAAGATGTTGAAGCCGCGCATCCGGC
GCTGGCGGACGACGACGTCCGTGTGCGTGTAGCTGAAGACGTGGCCGATGTGCAGCGAGCCCG
```

AGGCGGTCGGCGGCGGGGTGTCGACGACGAAGGTCTCCTCGCGGGGGCGCGACGGGTCGTATC
GGTACGTCCCGTCGGCCTCCCACAGGTCGGCCAGGCGCAGCTCGGCGACGGGCGAGTCGAAGT
GCTTCGGGAGCGTCGCGGGATCGATGGAGCGGAACGTCTTCTTGATCGTCACGTGGTCACCTG
CAGAACAGACCCCGCAGGAACCGCCCGCGGGGCCGGCATCCTACGTCGTCCCCCGGGTGCCGC
TCAAGGCGCGCCGCGCCCGCGCGGCGGCGATCCGCGATCGCATCCGCGCATCCGCCAGAGCCC
GGCGGCTCCGCCGGCGCGCGCGCGCCGTCCGTGGAGCCGAGAGGAGAGGCCGGCGCCCAGGTC
GTGGAGGACGCCGGCGGCGCCGCCGCGGAGATCGCGGAGAGGCGGGCGCATCGATCGCGGCGA
GGCCGGGGGCTCAGTCGTAGCGCTCGACGTGGACGTGCTTGCGGTGGACGCCGAGCTCGCCGC
GGGCGAGCTCGCGGACGGACGAGACCATCCGATCCAGGCCGCAGATGAAGACGTGCGGCGCCG
GATCTCCGCTCTTCTCCGCGAGCTCCCGGTAGAGCTCGGGCACGTGCGCCTGCACGTAGCCGC
GGCGGCCGGCCCACGACGGGCCGCCGCGCGAGAGCGTGATCTCGTAGCGGATCCGGTCGGATC
CGCGCGCGAGCGCCTCGAGCTCGTCGCGGTAGATGACGTCCTCCTCGAAGCGCGCGCCGAACA
GGATCCACAGGTGGGGCGCGGCCAGCCCCGCGCGCAGGGAGGCGCGCAGCATGCTCCGGAGCG
GCGTGATGCCGGTGCCGGTCGCGACGAACAAGGAGGGCGCGGAATCCCCGGGATCGCGGGTGA
AGAGCCCGTGCGGGCCGATGGCGCGGAGCGTGGCGCCGGGCTCGAGCCGGTGCAGGTGCTCCG
AGCCCGCCCCGCCCTGCACGAGCGTGACCGCGAGATCGAAGCGGGGCGAGCCGTCGGGCGCGG
ATGCGATGGAGTAGGCGCGCTTCACCTCGCCGCCCGGGAGCGGGAGGACGAGGTTGACCCACT
GGCCCGCCTCGAACAGAAACGACCTCCCGTCGGCGCGCTCGAACGAGAGCTCGCGCACGAAAG
GGCTGAGGGGCCGGGCGGCGACGAGGCGGGCTTCGAACGGTTCGGCGTGGATCATGGTCGGGG
CCCGGCGGGGCTCGGCTGCGAGGCCGCGCGGGTGGCGAGGTCTTACCGCAGCCTGCGCCCCGG
CCCAATCGCGATCGCCGCGGGAAGGGCGCCGCCGGAGGGCGCGCAATCGCGGGAATCACGGGC
TTCCGCCCCGTGCGCCGCCGGAGCGCGCGGCCGGGCCGCCGGCCCGCGCTCCGGCGGGGAGCC
GTCGCGGGCTCTACCGCACGCCCATGCGGCGGCGCTGCGGGATGTTCACCGCCGGCCGGGAGC
GATCCTGGTTGGGGAGCGCGCGCGGCGGGCGGGGATCCCGGTGCGCGGGCTTCTGCGCGGGGA
GCTGCCCTCGCTGAGCCGGGCGCTGGTCGGGCGACTTGGCCGAGCCCAGCGCGAGATCGGAGA
CGGGGAGATGCGCGCGTCGCTGCATAGAATCCTCCATGGAATCGGTCATCAACACATCGGGAA
GAGCACCCAGGCTGAAAGAAACCTTCGAAGAACCGGCTCTCATACACCCTCCATTCATCGTGC
GACCCCGGATTCAGGACGGATCGAACCCGCGAGGGACGCTGGCTCTCTGGGCCTCTCCCTGCT
CGCTCGACCGGCGCCCTCTCGACGCAACTCCGCCGTTCGTCGGACGGACGGTCCGCCTCGC
CGCACGCTCCCCGTCGAGACGACTCAGCGTCTCGACGTCAGGAGAGATGACGACTCGGCCCGT

```
CGCGCCACGACCCTTCCGGCTCGGTGCTTCGAGCGCGCGGCCAGCGAGCGAGGGGCGATCGCC
AGGAGATCACGAATCTCCCGGCCATCGGCCTCCAGCGCCTCGGGCTCGTTCGCTCGTCGCCCC
GCTCCGTCCCCGCGCGCGCACGACGCGAGCTCGCGCGGGGAACCGCGGGCCGCTGTCGTGGCT
GCTGATGCGCGACGATACAGGGGGGACGCCGTGCCTACCTGGGCAACAGGCGCTCATCTTCTA
CCACGGCGAGCACTACGGTGAGTGCTGCCATGAGTAGGCCCCTGAGGGTCCGCGCGACGGAGC
GTGGTGTCAGCGAGAGATGCGCATGGTGGACGCGGGCTACGCGTCGAGAGGGACACTAGCACT
CGACCTCGATCCTGCCCAGCACTTTTTGTCGGGGAGGGCTGCCCTCCCGCTGGCCGCTGGCCG
CTGGCCGCTCGCCGCTGGCCGCTCGCCGCTGGCCGCTGGCCGCTCGCCGCTGGCCGCTGGCCA
TGTGCGACGTGAGCTCGAGCAGCCCGCGGCTGACGGACAGACCCCGGAGTTCATCGAGCCGGT
GATGCCGAACCCGCCAAGCGAAAAAACGTATCCGTTCGGCAGGTCGTGGCCTATCATGCAAGC
TGCTCGATGCGCTGACAGGCTTCTTCGAGATCCTCGTCGGTCTTTGCGAAGCAAAACCGCATG
AAGCGACTCCCCTGCGTCCCTTCAAAGAAGGCGTCGCCTGGCACGCCCGCCACCCCGGTCTCG
TCGAGCAAGTAGATGGCTCGCTCTCGACCTGTCCTCCCGGGTAGGCGAGACACATCCGCCAGC
ACGTAGTACGTCCCCTGCGGCACGCAGGGTGGCAAGCCCGCTTTCTCCAGCGCCCGACAGAAC
CGGTCTCGCTTCCGTTCATATCCCTGGGCAAGCCCCGTGTAAAACGAGCGAGGAAGGCCGCGG
ATCCCGGCAGCGACTCCATGCTGCAGCGGCGTCGGCGCGCAGACATACAGCAGGTCGCTCATG
GCTCCAATGGCCTTCGCCCACCTGGCATCGGCCACGCTGTAGCCGATCCGCCATCCTGTGATG
CTGAAGGTCTTCGAGTAGCCGCCTATCGTGATCGTACGCTCGGACATGCGCGGAAGGGAGGCG
ACGCTGACGTGCTCACGGCCGTCGAAGATAAAGTACTCGTAAATTTCGTCCGTGATCACCATG
AGGTCATGGTGGCAGGCGAGATCGGCGATCTGTTCCAGCTCCATTCGGCCGAACACCTTCCCG
GAAGGATTTCCAGGAGAGTTCACCACGATCGCCTTGGTCTTCGGGGTGATCGCGCGCTCCAGC
TCGTCGCCGTCGACATTCCAGCTCAGGGATCGCGCCGTCACATACCGCGGAACAGCCTCGACG
GCGAGGATAGCCTGGGCGTGATAGGCATAAAACGGCTCGAAGAGCAGCACTTCGTCCCCAGGA
TTGAGCAAGGCCATGCAAGTGGCCTGAAAGGCCCCTGTCGCTCCGGCGCTCACCGTGATGTCA
GTCTCCGGATCCGCCGCGATGCCATTATGGCGAGCCAGCTTCGCCGCGATCGCATGGCGCAGC
TCCACGATGCCGTCGAAGCGCGAATATGTATTGCACCCCCGATCCATCGCCTCCTTCACCGCT
TGAAGGATCACCGAAGGAACTGGGGTATCACAGACGCCCTGGGACATATTGATCCCATGGACC
TTGGCGCACGCCAGGGTCATGGTACGGATATCGGACTGGGCGAGGCGAGCCGCACGATCACTC
GGTAGACTCTTCATCAGCGTGCTCCTGCTTCTGTTCTGCGGCTCTGCATGGTGTCTTCGGGTG
GGCTTGTCAGCTCGACGCGCCCATGCAGCGGCGCAGCCCTAGCGGCCGCAGGTCTGTCCACAC
```

TTCTTTGATGAAAGCGAGACATTCGGCTTTCGTGCCCTGTTTGCCCGCAGCCCTCCAGCCCCC
AGGTACGGGCTTGTCGGCGGGCCAGATCGAGTACTGCTCTTCGCCGTTCACCACGACCTGGCA
ACGCGTCTTGCTTTCGTCGTCCCGATTCATGATTTTCCTCGCCCTTCGTCAGCGCTGCGCGAG
CATGAAACGAATCGCTCATCGGCGCACAGGCGCGCGCCGGCTGCCCGGAGGCACTCCCACGCC
TCCCTCACGGCAACCTCATCGCTCCGGATGTTCCCGATGGCGACTCGGATCGTGTACCTGCCG
TGGAGACGGGTATGGGACAAAAATACCCTGCCCGACTTGTTGACCTCGTCCAGCAGCGCCTCG
TTGAGGCGATCGAGCTCGCGTTCGATCGACTCTCTCTCTGCCTCGTCCGCCGACCGCATGATG
CAAGCGAGCGCGGAGGGCCTCATGCGAAAGCAGACCGTACTGAACGGCGTCGGCGCGAGGCGC
TCCCAATCGGGATCGGCGTCCACCCACTGGGCGAGCTGCTGCCCCAATCGGAGGTGCTCCCGG
ATCCGGGCCGCCAGCCCTTCATGCCCGAAGTAGCGCACGATCATCCAGAGCTTCAGCGCTCGG
AAGCGCCGACCGAGCTGGATACCCCAGTCCATGTAATTCGTGACGTCGCCCTCGGTGCGGAGG
TATTCGGGCACCAGACTGAACGCGCGCTTCAGTCGGTCGGCGTCACGCACGTAGAGCACGCTG
CAATCCATGGGGGTGAACAGCCACTTGTGAGGGTTCACTACCAGCGAGTCCGCCCCCTCGCAG
CCCGCGAGCACGTCCCTGTGCTCGGGGACGATCGCGGCCATCCCCGCGTAGGCCGCGTCCACG
TGAAGCCATAGCCCGTGCTCCCGGCAAACGCTGACGATGGCGGGGATGGGGTCGACGCTCGTC
GTGGACGTCGTGCCCACCGTCGCCGCGACGCAGAAGGGTCGGAGGCCGGCCCCGAGGTCCTCC
ACGACGGCGGCGCGCAGCGCCTCGGGGACCATGCGGAAGGCCGGATCCGTGGGGATCTTCCGC
ACCCCCTCCTGCCCGATGCCGAGGGTGATGGCTGCCTTCTCGATGGATGAGTGCGCCTGCTCC
GACGCGTAGAGTCGCATGCGCCGCTGTCCCGCCATGCCCCGGAGCCGGATGGTCGGCTCGGCC
GAGTCGCGCGCGGCCGCGATCGCGACCATGCTGGCGGTCGACGCGGTGTCCATGATCGCGCCG
TGCAAGCCGGCGTCGAGATCCAGCATCTGACGCAGCCAGGAGAGGACGAGCTCCTCGAGCTCG
GTGGCCGCCGGCGACGTGCGCCATAGCATCACGTTGACGTTGAGGCACGCCGCGAGCAGCTCG
CCGAGGATCCCAGGACCAGACGCCGTGTTCGCGAAATACGCGAAGAATCGCGGATGATTCCAG
TGCGTGATCCCCGGCAGAATGATCTGCTCGAAATCGGTGAGCACGGCGTCCATCGGCTCCGGC
TCGACGGGCGGGGTCGGGGCCAGCCTGCCCTTCACGTCGCCGGGGCGGATCGCGGGAAAGACG
GGGTATCGATCCGGGTGGCCGAGGTAATCGGCCGCCCAATCGATGATTCTCATACCGATCCGG
CGGAACTCCTCCAGATCCATGTCCCCGAGCCGTTCTTTCCGCGGGTCGCTCACGTCAACCTCC
TCGCCCTGCCAGGACAGGATCCTCGAGGTCCCCTGGCTCCGGCGGTGGAAAGCGCTCCTTGAA
CGTGAAGGCCCACGGGGTCGGTCCGTAGCGCCGCAGGTGCTCGAGCCGATCCTGCCCCTCGCG
GACGGACGGGATGTGCCCGGCCGGGACCCACCACAGCACGAGGTAATGCGGCTCGAGATGCTC

GAACCACCGAGCGCGCTGTCGCAGGAACGCGGCATGATCCGCGGTGTAGGTGAAGGCGAACAG
GTGCTCGATGGAGGTCCATACCGACAGGGTCACGAGGAGCCGCTGGTCCGGGTACGGACGGAT
GGACACAGAGTTCCCCTCGGCCGTCTGCAGGCGCCACACGAACCCCTCGCTCCGATCGGCCAG
ATGGTTGATATGGTCGAGCCCCTGGACGAAGCCCTCCATGATCGGATCCTCCAGCGGAGCGCG
AATACATGCGAAGTTGTATTGCGCGATGTGGTGCCGATGCTCCGACATGTCGCTTTCCATCTC
CAGCTCCCGCTCACCAATCCCAGCGCTGCTCCGGGGAGCTCATCAGGGCAGACGCGACATCGA
TCCCGAAGCTCCGCCGCATCCCCTCGACGAAGGCGGCCTGGACCGCTTCGGCGACGGATCGGC
CTGCCTCCGGCAAGACCTCGGAGACAAAGAAGAACCGCCTCGTGGAAGGGACAATCTTGCCCC
GCTCCGCCTGGCGCCATACGAAGTGCCTCGTCACCAGTCCCTCCGCGTCGGCATACCCGACCT
CGCCGGCGCCGACCGCCACGCTGCCGCCTGAGCCGAGCTCCACGAACGCCTCACCGCCTCGCG
AGATCTCGAGGCGAACGTCCGGGCCAGCCAGATCGCCGAGGTCCCAAGCGCCGACGGGGACGG
CGAACCGCAGCGACAGGAGGTTGTAAAAATCGACGAATGCGTTGATGTGCGGCAGCTCTCCAC
CACCGAGGACCCGCTTCGCCAGCGCCTCGATCGAGCTCGGAAATTTCTTGCCAGAGACCCCCA
CTCGCTTCATCGCCTCGCGCCAGGCAGCCACGTGCGGATGCGACTGGGCGTTTTCGTGGCCCC
AGCTCCGTCGCAGCTCCTCCTCGACCTTCCGGAGCTCCTCCAGCACGGCCGGCCGCTCTGCGG
CGTTGTCCAGGCCTTCCCCGTACCCGGTGACCAAGATCATCCCAGGAAACGACTCCCAGATTC
GCGGATCGACGATGAATGCCATGTGCCTCCTGCCCCTCGAGAGCGATCGCCTCGATCGACACC
AGGCTGTGGATGCATGAGCCGGGCCGTGCGGACGCAGGACCCCGCTACTCATGGCTCTTCGTG
GCCGATGAACAGGTCCTCCACCCGTCGATCGTGCTCGGTGCCCCGATCCGTCCAGTCCCACCC
GCCGGCGACCGCGATGTTTGCACCCGAGACGTACGAGGCGCGGTCGGAGGCGAGGAACGCCAC
AGCATCTGCGACCTCGCTGGCGCGCCCCAGGCGGCCCATGGGGACGCGCCGCTCCATCCACTC
CTTCTGCGCGGGCGGAAGGTATCCGTTGTCGATGAGCCCTGGAGACACACAGTTGACCAGGAT
TCCATGAGGCGCCTCCTCCGTGGCCAGGCTGCGCGTGAGGATGAGCACGCCGGTCTTCGCGAT
CGAGTACGCCGCCACGTTCGGCGCGCCGCGGATCGCGTACGTGGGGCTCAACCCGATATTGAT
GATCCGGCCGCTCTTTCGCTGGCGCATGCGCGCCACGGCCGCGCGACAGAGGTAATGAACGCT
GCTCAGGTTGCTGTCCATGACGTTGCGCCATTCGTCGTCGGTCATCGCCGCAAGCGGCTTGAA
GAAGAAGTCGCCCACGTTATTGACGAGGATGTCGATGGGGCCCAGCTGCGCCTCGACGCTGGA
GAAGAGCTCCGCGGCCGCGTTGGGGCGGGTGACGTCGGCCTGCACCACCATGGTTCGTCGCCC
GAGCGCGCGGATCTCGGCCGCCGTCTGCTCGGCCGCATCCTTGTTCGAATGGTAATTGACGGC
GACGTCCGCGCCTTGCTCCGCGAGGCGCAGCGCGATCGCCTTGCCAATTCCGCGCGAGCTACC

```
GGTGACCAGGGCGACGCGCCCGGCGAGCTCCAGCGATCGCGCCTGTGGCAGGGCCGGAGCAGC
CTCCTGGTGGAGCTCGACGTCGACGGGGAGCTCCACGTGGTAGCTCGTCTCTCGCGGAGCCGC
GCAGTACCTCTCGTAGAACGCCTCGAGGACGGGCTCTTCGCGTCGCATGATGTCCGCGTGGGA
TTCGGCGCTGCGCCACGGATAGAGCACCAGGATCTCGTCGGGGCGCACGGTGCTCTGAAAGAA
TCGCGCATGACCGCACCACCCCGGGTGTTCGTGGGCTCCTGGGCCGAGCAGATCATCCATTTT
TTGCATGATCTGCGTGGCCTCGCCCTCCATGCCGGGTTTGATGCGCCATCGCTCCATAACGAG
GATCATGTCTTGGCTCCTGTTCGTCATCGCCGTTTCGATCTGGGGGGCTGCCCGCGCTCTCG
AGGGCGCGCCCCTTGTATTGGCCGCGGATGGTCTGGGTAGCGCTCGCGAGCTTTCGCTTGTGG
GCGGCGTTCAGGCTTGCGCCTTGATTGACGAACCGCTCGCAGACGAATGCGTGCGATTCATAT
GCGGTCGCGAGCGCCCACAGGTAGAGACGGTGGCTACGAGCGAGCTTGGGCGGAAGCGCCCTC
AAGGCGGGTCGGATGGATTTGAGGGTGCGAATCAAGCGGGCGTGCTCTTCGAAGAAGAGACCG
CTGAATCCATCCTGGACGAACGGCGGGGCCATGGTAGGTCGCACGATGTTGTTGTAATCATCA
TTCGAGAAATCGCCGGTAAACCGGAAAGATGCCGCCGTTGCCCACCAGAGCTGCGTGGAGATC
GCGAGCGCCTCTCGCGCCCTCGGAACGTCACGGTTGGCGAGCGCCTCCGACAGCCACCTGTGC
GCCATGAGCAGCGCCTGGACCAGCACGAAGAACGCGTGATGCCCCAATACCCAGCGGCCCAGC
GCGCCGTCCGGAGCGCCCGCCTCCGCCGGAGGCGCAGCCAGCTGCGGGATGCCATTCCATGAT
TTTGGCCTTCGCTTGCCGGAGAACTGGTGGAGGATGTCCTCGATGGACGCACAGAGATGCCCC
ATTTCCATGGGTTGCAGGGAAGTACCTTTCAGGCTTTCGCGGATCATTCGGTAATATGCGACG
ATCACCGCTTCGCAGTACGTCTCGAGCGGGTCTGGGGCGGAGACCCGGTGCACATGGAAATAG
GCGTCGTACTCCGCCTCCCGGTCCGACAGGCTCCCCGGGATGTCCGGATCTGCCGCTTGCGCC
TGCCATCGCTCGAGGATGGGGACGGCGAGGGGGGCGAGATGCTCGGCGAGCACCGCCAGGTCG
CCCTCCGGCGCATGCGCCAACAAGACCTCGAAGGCCTCGGCTGCGACCTTGGAGGTCTCGCCC
ATTCCGACGGCCTCGATGGCTTGGGGCAGCGGTAGATGGATGGTATATTTAGCCATGATTTGC
CCGAAGATTGCCGCTGCGTCGACAGATCTTTCGCGAGCCGGAACGCCATTTCCACTGCTCTGG
CTCTCAATATTGAATTGAGCCCTGGCGACTGCCATAGGCCCAGTCGCTCGACACAGTGTACGG
AGCGGCCCGATGCTTTCTCCTTTTTTAGTCCTGCACCGAATACTTCTGTTGGGCGCCAAAGAT
CCCTTGCCGAGACTGTCCGGCGAGATGTCGTGTGCGAAGCGTCCGCACGTCCAGCGGGCCCAT
GCGTTGCTAGAGCATAAAACGGTTCGATGCCTGGTCGAGAGGGAGACGCGAGGAGCCTCCCTT
TGGGACGGATGAGGAATTTCGTGACCGAAATGTCGGCAGGAACAGCGGCGCAGAAGCGGCGCA
TCGATGGGGAACCATGGGTTACGAAGACATTGATGATAATGTCGACGCAATCGCAATCGTCGC
```

```
GATGAGCGGCCGCTTCCCCGGCGCGAGAAACGTCGAGGAGCTGTGGCAGAAGCTCCGCGCTGG
CGTGGAATGCGTCGTCACCTTCACAGAGGCCGAGGCGCTCGCCGCGGGGTGAGCCGCGAGAT
GCTCGCGAATCCCAGCTACGTGCGCAGAGGCGCGCCGCTCGACGGCGTGGAGCTCTTCGACGC
CTCGTTCTTCGGGTTCAGCCCGAGGGAGGCAGAGAGCATGGATCCGCAGCAGCGCATCTTCCT
GGAGGTCGCCTGGGAGGCCCTCGAGCGCGCCGGTTACGACCCCGATGCCCATTCCGGGCCTAT
CGGCGTCTTCGCGGGCAGCGCCCCGAGCGGCTACCACTCCCTGGCGCAGTCCGACCCGGAGAT
CCTAGGCGCCCTCGGCCACTACCAACTGACGCTGAACAACGACAAGGATTATCTCACCACACA
CGCCTCGTACAAGCTCAATCTGCGGGGCCCGAGCGTGTGCGTGCAGACGTCCTGCTCGACCTC
GCTCGTGGCCGTGGTCATGGCCTGCCAGAGCCTGCTCAACCACGAGTGCGACATGGCGCTCGC
GGGTGGCGTGGGGATCCATGCGCATCAGCGGAGGGGCTATCTGTATCAGGAGAACGGCATCTC
TTCGCCCGATGGGCATTGCCGCGCCTTCGATGTGGCCGCCAAGGGCACCGTGGGCGGCAGTGG
CATAGGCATCGTCGTCCTGAAGCGGCTCGCCGACGCGCTCGCCGACGGCGACCACGTGCACGC
GGTGATTCGAGGAGCGGCGATCAACAACGACGGCTCGAGCAAGATCGGTTACACCGCGCCGAG
CGTGCAGGGGCAGGCCGAGGTGATCGGCATGGCCCAGGCGCTCGCCGGCGTGGAGCCGGATGA
CATCAGCTACATCGAGGCGCACGGCACGGGGACGCCGCTCGGCGATCCCATCGAGATCGCAGC
CCTCACGCGCGTGTTCCGGGCGAAGACCGCACGAAGGCAGTTCTGCGCCATCGGCTCGCTCAA
GACCAACCTCGGCCACCTCGATGCCGCCGCGGGCGTCGCCTCGCTGATCAAAACGGTCATGGC
CCTCGAGCACCGCGAGCTGCCCCCGAGCCTGCACTTCGAGCGTCCGAATCCGAAGCTCGAGCT
GGAGAGCAGCCCTTTCTACGTCAACACCCGCCTCACTCCGTGGCACGCGGCACGAGGTCCGCG
CCGCGCTGGCGTCAGCTCGTTCGGCATCGGCGGCACCAACGCGCACGTGGTCCTCGAAGAAGC
TCCGGCCCCGCCTCCGAGCGGCCCCTCGCGGCGTTGGCAGCTCCTCACCCTCGCGGCTCGCTC
CGAGGCCGGGCTCGCGCGGGCCACGGCCGACATGATCGAGCACCTCGATCGCCACTCCGGCAC
ATCGATCGCCGATGTCACGTACACGAGCCACGTGGGGCGCCGGGCCTGGCCCTTCCGGCGAGC
GGTCGTCGGCGAGAGCGCCGCGGATCTCCGCGCCGCGCTCGCGAGCGAGGGCTCGCCGCGCTC
GATCTCGTCATGCCAGGCGGCGAGGGAGAGGCCCGTCGTCTTCCTGTTCCCCGGTCAGGGAGC
GCAGCACCTCTTCATGGCGCGGGAGCTGTACGAGGTCGAGCCGATCTTCCGGCAGTCCCTCGA
CCGCTGCGCCGAGCTCCTGCGCGGCCCGCTCGGCCTCGATCTGCGGCAGGTCCTCTACCCCGC
CGAGGGGCAGCGCGACGACGCCGAGCAGGAGCTCGGTAGGACCGCGATCGCCCAGCCCGCGCT
GTTCGCCATCGAGCTCTCGCTCGCCAAGCTGTGGATGGCCTGGGGGATCGTCCCCCAGGCGAT
GATCGGCCACAGCGTCGGCGAGTTCGCCGCGGCTTGTCTGGCGGGCATCTTCCGCGAAGAGGA
```

```
CGCGCTCCGCCTCGTCGCCGAGCGGGGCCGCCTGATGCAACAGATGCCGCCCGGCGCGATGCT
GGCGGTGCCCCTCGCGGAGCCCGAGCTCGCCCCCTACCTCAGCGACGACATCTCGCTCGCGGC
GATCAACGGTCCGGCTCTCTCGGTGGTCGCTGGGCCGATCGAGGCCATCGACGCGCTCGCGGC
CGAGCTCTTGGACCACGGGCTCTCGTGCCGGCGACTCCACACGCGGCACGCCTTCCACTCGAA
GATGATGGCCCCCGTCGTTGACGCCTTTACCCGATGCGTGTCCGCGGTCGAGCGCCGCCCGCC
GTCAGGCCACTTCCTCTCGACCCTGACGGGCGGCTGGATCTCCCCCGAAGCAGCGACCATCCC
CGCATACTGGGCCCGGCAGCTCGTGGAGCCGGTGCGCTTCGCCCAGGCCGTGAGGCAGCTGCT
GTCCGAGTCGACGTGGCTCTGGCTCGAGCTGGGTCCGGGCCAGACCCTGAGCCCGCTCGTACG
GCAGCAGGCCCGCGCGGATGGCGGCCAGGTGGTCGTCGCCTCGCTGCCGCGCGCGAAGGACGC
GGGCGCCGACCACCTCGCGGTCATCGAGGCGCTCGGCCGTGTCTGGAGCGCTGGTGGGACGGT
CGACTGGAAGCGCTTTCACGAGGGCGAGGCGCGGCGGCGGGTGCTGCTACCGACCTACCCCTT
CGAGCGGCAACGATACTGGGCCTCTCCGCGCCACACGAGCGCTCCGCCGGAAGCGATAATCAA
GCCGCTCCTCGCGAAGAACCCAAACGTCGCCGATTGGTTCTTCCTCCCTGCCTGGCGGCGCTC
GGATCCTCCGGTCTCGTTCGACGCGCAGGCGGTGACCACGCGGCGCTCTACGTGGCTCGTCTT
CATCGGGGACGAGGGCCTCGGCGCGGCGCTGGTGGAGGGCCTCGCGCGGCGGGGGCACGAGGT
CGTCGCGGTGGTCACGGGTGAGAGGTTCGAGCAGACGGGCACGCAGCGCTACACGATCGATCC
CGCCGCGAATGGCGATGTTGCGTCCCTCTTCGCGCGGCTCGAAATCGAAGGGCGCATGCCGGA
CCGGATCGTCCATGCCTTCTGCACGTCGCCTGCGGACGGCGCGCGCATCGAGCGCGGAGCCGC
GCTGGAGATCGAGCGCAGGCTGGGCTTCGATAGCCTCCTCCTCCTCGCCCAGGTGATCGCCGC
ACAAAGGCATCCGAAGCCGCTGATGCTCGGCGTGATCACGACCCGGGCGCACTCCGTCATCGG
AACCGAGATCATCGAGCCCCTGCGCGCTCTGGTGCTCGGCCCCTGCCGCGTCATCCCGCAAGA
AATACCCCATGTCTCGTGCCGGAACATCGATATCGATCTCCCGGGCGAAGGCGGGCGCGCGGA
GATCGCGGCGCGCCTGATCGCCGATCTGGAGCGAGAGTCGCCCGACTCGGTGGTGGCCTACCG
CGGCGGCCGGCGCTGGGTCGAGAGCATAGAGCTCACCGATGTCGGCCGGCGGTCAGCTGGCGC
CGCCCCGCGCCTCCGCCAGCGCGGGGCGTACCTCATTACCGGCGGCCTGGGGGGCATCGGCCT
CGTGGCTGCAGAGCTCTTGGCCCGAGAGGCGCACGCACGGCTGATCCTGGTTGGGCGGACAGG
CCTGCCAGCGCGGCAGGGGTGGGACGACTGGCTCGCGGCGCACGGCGCGGGCGACGCGACGAG
CCGAAAGATCCTCCGGATCCGCGCGCTCGAGGAGGCCGGCGCCGAGGTGAAGATCGCCGCGGC
CGACGTCTCCGATTTCAATGCGATGCGGAGCGTCATCGAGGAGGCCCGGACGCGCTTCGGCCG
CATCGACGGCGTCATTCACTCCGCCGGCATCGCGAGTGGAGGCATGATCCAGCTCAGGACGCC
```

GATGGCGGCTTGGCGCGTGATGGCGCCGAAGGTCGGCGGCACGCTCGTGCTCGATGCGCTCCT
CCGGGACGAGCGTCCCGACTTCCTCCTGATCTGCTCGTCGTTGGCCTCGCTGGTCGGCGGCGC
CACCCAGATCGATTACTGCGCCGCCAACGCCTTCCTCGACGCCTACGCGCAGAGCCGCGAGGG
CGAGGAGGGATGCCGCGTCATCTCGGTGCAATGGACACGTGGAGTGACGTCGGGATGGCGGT
GGACTTCAAGCTCCCGGCCGATCTCCAAGAGGGGCGCCGCGAGAGCCTGAAGCGGGGCATCAG
CTCGAGCGAGGGCGCCGAGGTGCTCGGCCGCATCTTGAGCGCAGGCATGAGCGGCCCGCTGGC
GATTTGCACGTCGGATCTACCAGCGTACAAGCAGTCTGTCACGACACGCCGATCGCAGCACGA
GCAAACTCCCGCCGCCCGGCCGATGCACTCGCGCCCAACGACCACGGGAGCCTATGTCGCTCC
CGAGACCGAGACCGAACGGCGCATCGCCGCGATCTGGCAGGATCTCCTCGGCCTCGAGCAGGT
AGGCGCAAACGACGATTTCCTCCAGCTGGGCGGCCATTCGCTGTTGGCCACGCAGGTCCTGTC
TCGCGTCCTGCAGACCCTCAAGGTGGGGATCTCGTTGCCGCAGTTCTTCGATGCGCCGACGGT
CGCAGGGCTTTCGCGCCTGGTCGACGCAGCACGGGCCGAAGGCGCCGGACCCGTCGCGCCGGC
AATCGGCCGTGTCGAGCGAGACGCCTACCGAATCAAGCCGCCCGCGGCCGAACAGGCCGCCCG
CACCAAGCCGTAACAAGAAGGGGATCGAGTCATGGAACCCGTCGGCGGCGTGGACATGAATCA
GCCCGCAAAGCAGCAGGAGACCTGCGTCTTCCCGACCTCCTTCGCGCAGCGGCGGCTCTGGTT
CCTCGACCAGCTCGAGCCGGGGAGCGCCGTCTACAACATGCCCGCCTCCTTCCGGACGCGCGG
GCCGTACGACGTCGACTCGCTCGTGCGCAGCGTGAACGAGATCGTGCGGCGCCACGAGTCGCT
GCGCACGACCGTCGATGTCATCGATGGCGAACCCGTGCAGGTGATCGCCCCCTCGCTGCGCAT
CGAGGTGCCCGTCGTGGACCTGAGCGAGATCGACGAGCCGGAGCGAGAGGCGGAGGCCCGGCG
GCTCATGGCGGAGGAGAGCCGCCGCCCCTTCGATCTCACGCGAGGGCCGCTGCTCCGAGCCAA
GCTGCTCCGGCTCGGCGAGGCCGATCACGTGCTGATCTTGACGATGCATCATATCGTCTCCGA
CGGCTGGTCGATGGACGTGCTGTTCAAGGAGCTTTCCACGCTCTACGCCGCCTTCCACGAGGG
CCGCCCGTCGCCGCTCCCGGAGCTGCCGATTCAATACGCCGACTTCGCGGTGTGGCAGCGGGA
GCTGCTCCAGGGCGAAGTTCTGGAATCGCACCTCGGGTACTGGAGAGAGCACCTCCGCGGCGC
CCCCACGCTGCTGGAGCTTCCGATGGACCGGCCCCGGCCGCCGGCGCAGACGTTCCGGGGCTC
CCAGCGCGCGTTCCGACTCCCACTCTCCCTGCAACAGGCGGTGCAGGCGCTCAGCCGGCAGGA
AGGCGCGACCCCCTTCATGACGCTGCTGACGGCGTTCAGCGTGCTGCTCTCGCGTTATGCGCG
GCAGAGCGATCTGGTGGTTGGCACGCCCATCGCGAATCGCACCCGAGCAGAGCTGGAGGGGCT
GATCGGCTTCTTCGTCAACATGCTGGCGCTGCGCATCGACCTCGGGGGCGACCCGAGCTTCCG
CGAGCTGCTCGGGCGGGTGCGGGAGGTGACGTTGGGCGCCTACGCGCACCAGGACCTGCCCTT

CGAACGGCTGGTGGAGGAGCTGTCACCAGGGCGGAGCCCCAGCCACAGCCCCTTGTTCCAGGT
GTCCTTCACGTTGCAGAACACCCCGATGGATGCGACGAACAGAGCAGACATTGCATCGGGTGG
CGCGCCGCTGGTGGAAATGAAGGCGGCGAAATTCGATCTGATCCTGGAGCTCTCGGAATCGCC
GCAAGGGTTGCTCGGCACGTTCGAGTACAACACCGACCTGTTCGACGCCGGCACCATCGAGCG
GATGGCCGGCCACCTGGAGGTGCTGCTCTCCAGCGCCGTCGCGGCGCCGGATCGACCCATTGC
GGAGCTGCCGCTCATGGGGGCCGAGGAGCGCAGTCGGGTATTGGTGGAGTGGAACTCCACTGC
CGCGCTGTATCCCGAGGACCATTGCATGCACGAGCTGTTCGAGCAGCAAGTGGAGCGGTCGCC
CGAGGCGACCGCGGTGCTCCTCCAGCAGCAGACGTTGACGTATCGAGAGCTGAACATGCGCGC
CAATCAGCTCGCGCATCACCTGCGGAGCCTGGGCGTGGGCCCAGAGGTGCGCGTCGGGTTGTA
TCTCGAACGGTCAATCGAGACGGTCGTGGCGATCCTCGGCGTGCTCAAGGCTGGCGGGGCCTA
CGTGCCGCTCGATCCGACGTACCCCAGCGAGCGCCTCGGGCTCATGATGGCGGACGCAGCGCC
CTCGGTGCTGCTCACGCAGGCGTCGCTCCTCTCGAAGCTGCCGCCCCACGGGGATGCAACGCT
GGTACAGCTCGACGCGCTGCACGAAGCGCTCTCCAGGCTGCCACACCATACCCCGCGGAGCGG
CGTCACCGCCCAGAACCTCGCATACGTCATGTACACTTCCGGCTCGACCGGGCGGCCCAAGGG
CGTGCTCGTCGAGCACCGCGGCCTCTGCAACCTGCCCACCGTGCAGGCCAAGCTCTATGGAAT
CGCGCCGGGCGACAGGCTCCTCCAGTTCGCGCCGCTCTGCTTCGACACATCGTTCTGCGAGAT
CGCGCTCGCGTTGCTCTCGGGAGCGACGCTGGTCATGGGCACGGCGGACGAGCTTCTCCCGGG
ACCTCCGCTGGTCGAGCTGCTGAAGAAGCACGCGGTCACGGCGATGCTCCTGGCCCCTACCGT
GCTCGCAGCGCTGCCAGAACAACAGAGCGCGGCGTTGCCGCTGCGCGTGCTCACGATGGCCGG
TGAGGCGTGCCCGGCGGAGCTCGTCAAGCGCTGGAAGGCACCCGGACGGCGCCTGTTCAACTC
CTATGGCCCGACCGAGACGACCATTTGGGCAAGCTCCGCAGCGGACCTGTCCGACGAACGGAT
CCCGCCCATCGGCCGTCCGATTGCCAATACGCAAATCTACGTGCTCGACGAAGCGCTCGAGCC
GGTGCCCATCGGCGTGCCGGGCGAGATCTTCATCGGCGGCGTGGGCGTCGCCCGGGGATATCA
CGGGCGTCCGGACCTGACGGCCGAGCGATTCGTACCCGACCCCTTCGGGCAAACCAAAGGGGC
GCGCCTGTATCGGACCGGCGATCGGGCGCGCTGGCTGCCGGACGGAAACCTCGAGTTTCTCGG
TCGAAACGACGAGCAGGTGAAGGTCCGCGGTGTCCGCATCGAGCTGGAGGAGATCCGCGCGGC
GTTGCTCAAGCACCCGGCGGTCGCTCAAGCCGTGGCCGTGGTGCGCGAGGACACGCCGGGGGA
CAAGCGGCTCGTCGCGTATGTCGTCGGACGCGGAGGAGCGCGCGTGACCGCCGCGGAGCTGCG
CCAGTCCGTGAGCGAGCGATTGCCTGCGACCATGGTGCCATCGTCCTTCGTGGCGCTCGACGC
CTTGCCCCTGACGCCGAATGGCAAGGTGGACCGCCGCGCGCTGCCGGAGCCCGAGCAGAGCGC

CGGCGGCGAGGACCACGTCGCGCCGCGCAACGCCGTCGAGGAGGAGCTCGCCAGGATCTGGGC
GAGCGTCCTCCGGCTCGAAAGGGTCGGCGTCCACGACAACTTCTTCGAGATCGGCGGCGACTC
GATCCTGAGCATCCAGATCGTGGTGCGCGCGCAGCAGGCAGGGCTGCGCCTCACCCCGCGTCA
GATGTTCCAGCACCAGACCATCGCCGAGCTTTCGACCGTGGCTAGAGCCGTCGAGGCGGTCCA
CGTCGAGCAGGACCCGGTGACCGGTCCCGCGCCGCTCACGCCGGTGCAGCGCTGGTGGCTGGA
GCAGGAGGCGGCCGAGCCGCACCACTTCAACCAGTCGATCTTCCTCGAGGTACGCGAGCGGCT
CGACGAGAGCGCGCTGGAGCAGGCCATCGCGCATCTGATCGACCACCACGACGCGCTCCGGTT
GCGCCTCGCGCGCGACGAACGCGGCGCCCACCAGGTCTTCGCCGCGCCGGGAGGCTCGACCCC
ATTTCAGCGCGTCGACCTCGGGGCGCTGCCCAGCGCCGAGCAGATCTCCGCCATGGAGAAGGC
CGCGAGCGAGGCGCAGGCGAGCCTCGATCTGGCCGCGGGCCCGGTCGTCCGCGCCGTGCTCTT
CGACCTCGGCGAGGTCGCCCCGCAACGGCTGCTCGTCATCGCCCACCATATTGCGGTCGACAG
CGTCTCCTGGCGGATCCTGCTCGACGATCTCTTTGGGGCCTATGAGCAGGCGCGCCGCGGCGA
GGCCGTACGCCTGCCGCCCAAGACCACGTCGGTCAAGCGCTGGGCCGAGCTGCTCACCGAGCA
CGCCGGCTCCGAGGCCGTCAAGGCGGAGCTCGGCTACTGGCTCGACTCATCGCGACGAACGGT
AGCTCCGCTGCCCGTGGATCGACGGGCCGGCGAGGACGTGTGGGGCTCGGCGCGCCACATCGT
CGTCTCGCTCACGCCGGAGCAGACGGAGCAGCTCCTGCGCGAGGTGCCGCAGGCGTACCGCAC
ACGGATCGACGACGCGCTCCTCACTGCGTTCGCGCAGGCCATCGCTCGGTGGACGGGCTCGCC
GGCGGTGCTCCTCGACCTCGAGGGTCACGGGCGCGAGGAGCTCGCCGGCGTAGACCTCACGCG
CACGGTCGGCTGGTTTACGGCCATGTACCCGATCCTACTCCGCGTCGACGCGGCGGATCCGGG
TGAGGCGCTCAAATCGATCAAGGAGCAGCTCCGCGCCGTGCCAGGCCGCGGGCTCGGCTACGG
CTTGTTGCGTTACCTTCGGTCCGATACCATCGCCGAGGTCCGCGCGTTGCCGCAGGCCGAGCT
CTGCTTCAACTACCTCGGCCAGCTCGATCAGGCGATCCCCGAGGCTGCACCGTTCCGGCCGGC
GCGCGAGTATCAAGGCTCGGAGCGCAGCCCCGGCGCCCATCGCGCCCACCTCATCGAGGTGAA
CGCGAGCATCGCCAATGGGCGCCTGTACGCCACGTGGACGTACAGCGAGCGCCGCCACGAGCC
CGAAACCATCGAGCGCGTCGCGGCGAGCTTCGTCACGGCGCTCCGCGCGCTCATCGCGCACTG
CACCTTGCCCGAGGTCGGCGGCAACACGCCTTCCGACTTCGACAAGGTGCGCCTGCGCCAGGA
GACCATCGATGCTCTCGACGCAATCGACGCGGGCCCCGGGCCGTCTGCGAGGGGGAGCCGAAT
CGAAGACGTCTACCCGCTCTCGCCGCTCCAGGAGGGCATCCTGTTCCACACGCTCTACGCCAC
CGATTACACGGCGTATGTCGAGCAGTTCCACTGGACGCTGGAGGGCGATTTCGACGCCGAGGC
GTTCACCCGCGCCCTCCAGGACGTCGTCGCTCGGCATGCCGCCCTGCGCACGTCGTTCGCCTG

GGAGCGCCTCGATGCTCCACTTCAGATCGTCCGCACGGGCGCGGTCCTCCCCGTCGAGCACCA
GGACCTACGCGGCCTCGCCGCGGAGGAGCAGACCGCGCACATCTCCCGTTACGTCGAGGCAGA
GCGCCAGCGCCGGTTCGATCTGCGAAAGGCGCCCCTCATGCGCGCCGGGCTGCTCCGGCTCCG
CAAGGACGCCTGGTGCCTCGTCGAGACCATCCACCACCTGATCCTGGACGGCTGGTCGACACA
AATCTTGCTCAAAGAAGTGTTCACGCTCTACGAGGCGCACCGCGGACACCGTGGGCATCTCGC
GCTGGAGCTCGAGCAGCCGCGGCCCTACGGCGATTACATCGGCTGGCTCGCGAAGCAGGACCA
GGTGCGCACCGCGGCCTTCTGGCGGCGCGAGCTCGAGGGCTTCTCCGCGCCGACGCCGCTCGG
CGTCGACCGCGCTGTGCCGCACGACGACGGCGGCCCGCGGTTTGGTTGGCGCCGCATCGCCCT
CTCGGGCGACGACGCGGCCCGGCTCGCCGCCTTCGCGCGTCAGCATCAGCTCACGATGAGCAC
GCTGGTGCAAGGCGCGTGGGCGCTGCTCTTGTCACGCTACAGCGGCGATCCCGACGTGCTCTT
CGGTATGACCGTCTCGGGCCGCTCGGCGCCGATTCCCGGTATCGAGCGCATGACCGGCCTCTT
CATCAACACCATTCCGGTGCGCGTGCGCGAGCCTGCCGACGCGTCGGTGCTCGCGTGGCTCAA
GGCGCTCCAGGAGCACGAGGCAGAGCTGCTCGAGCACGAGCACAGCCCGCTGGTCGAGGTCCA
GGCCCATAGCGACGTGCCGCGCGGACCCCGCTCTTCGAGAGCCTCGTCGTGTTCGAGAACTA
CCCGGTGCAGGTCATCTTCGAGGCCCCTCCGGTCGAGGGGCCGACGCGCGCGGAGGAGGGCCT
CCGCATGATCGATGCGCAGTATATCAGTGATCCACCGTATCCGCTGACGGTCGTCGCGGCCTT
CCATGGGACGCTTTATCTCAATATTGGCTACGAGCGCCGCCGGTTCGACGACCAGGCCGTCGA
ACGGATGATCGGGCACGTCACGACGCTGCTCCGGGGCTTCGTGCAGAGGCCCGAGACGTCGGT
CCGCGATCTGCCGTTGCTGACGGCCGAGGAGGAGCGCACCCAGCTCCACGCGTGGAATGCCAC
GGCCGCGCCGTATCCCGAGGGCCATTGCATGCACGAGCTGTTCGAGCAGCAAGTGGAGCGGTC
GCCCGAGGCGACCGCGGTGCTCCTCCAGCAGCAGACGTTGACGTATCGAGAGCTGAACATACG
CGCCAATCAGCTCGCGCATCACCTGCGGAGCCTCGGCGTGGGCCCAGAAGTGCGCGTGGGCTT
GTGTCTCGAACGGTCGATCGAGACGGTCGTGGCGATCCTCGGCGTGCTCAAGGCAGGCGGGGT
CTACGTGCCGCTCGACCCGACGTACCCCAGCGAGCGCCTCGGGCTCATGATGGAGGACGCGGC
GCCCTCGGTGCTGCTCACGCAGACGTCGCTCCTCTCGAAGCTGCCGCCCCACGGGGATGCAAC
GCTGGTACAGCTCGACGCGCTGCACGAAGCGCTCTCCAGGCTGCCACACCATACCCCGCGGAG
CGGCGTCACGGCCCAGAACCTCGCATACGTCATGTACACTTCCGGCTCGACCGGGCGGCCCAA
GGGCGTGCTCGTCGAGCACCGCGGCCTGTGCAATCTGCCCACCGTGCAGGCCAAGCTCTATGC
AATCGCGCCGAGCGACCGGCTCCTCCAGTTCGCGCCGCTCTGCTTCGACACATCGTTCTGCGA
GATCGCGCTCGCGTTGCTCTCGGGAGCGACGCTGGTGATGGGCACGGCGGACGAGCTCCTCCC

GGGACCTCCGCTGGTCGAGCTGCTGAAAAAGCACGCGGTCACGGCGATGCTCCTGGCCCCTTC
GGTGCTCGCAGCGCTGCCAGAACAACAGAGCGCGGCGTTGCCGCTGCGCGTGCTCGCGATGGC
CGGCGAGGCGTGCCCGGCGGAGCTCGTCAAGCGCTGGAAGGCACCCGGACGGCGCCTGTTCAA
CTCCTATGGCCCGACCGAGACCACCATTTGGCAAGCTCCGCAGCGGACCTGTCCGACGAACG
GATCCCGCCCATCGGCCGTCCGATTGCCAATACGCAAATCTACGTGCTCGACGAAGCGCTCGA
GCCGGTGCCCATCGGCGTGCCGGGCGAGATCTTCATCGGCGGCGTGGGCGTCGCCCGGGGATA
TCACGGGCGGCCGGACCTGACGGCCGAGCGATTCGTACCCGACCCCTTCGGGCAAACCAAAGG
GGCGCGCCTGTATCGGACCGGCGATCGGGCGCGCTGGCTGCCGGACGGCAACCTCGAGTTTCT
CGGTCGAAACGACGAGCAGGTGAAGGTCCGCGGTATCCGCATCGAGCTGGAGGAGATCCGCGC
GGCGTTGCTGAAGCACCCGGCGGTCGCTCAAGCCGTGGCCGTGGTGCGCGAGGACGCGCCGGG
GGACAAGCGGCTCGTCGCGTATGTCGTCGGACGCGGAGGAGCGCGCCTGACCGCCGCGGAGCT
GCGCCAGTCCGTGAGCGAGCGATTGCCCGCGACCATGGTGCCGTCGTCCTTCGTGGCGCTCGA
CGCCCTGCCCCTCACGCCGAACGGCAAGGTGGACCGCCGCGCGCTGCCGGAGCCCGAGCGGAG
CGCCGGCGGCGAGGACCACGTCGCACCGCGCAACGCCATCGAGGAGGAGCTCACACGAATCTG
GGCCGACGTACTTGGGGCAAAGCGGGTCGGTGTGCACGACAATTTCTTCGATCTCGGCGGCCA
TTCCCTGCTGCTCGTCCGGGTGCATGATCGGCTCGGCCAGCGGTTCGATCGGCCGCCCTCGAT
GGTCGACCTCTTCACCTATCCGACCGTGGCGTCGCTCGCGCGGTTCCTTGGCGAACGGGCGAA
CGGCAAGCAATCGCCGAGGGAGGCCGCGGCGGACGTCACGGAGCGCGGCCGGCGCCGCCTGGA
GGCGCGGGCGCGGCGGGCGAAGGCCATCCGTGGCCCGACCTGACCCGGGCACCCTTCCAAGCC
CCGCCGTTCCTCGCACATCCGCCGCCTCGAGCGCCGCGTCCAGCGCCGCCGTTCGCCGACGAG
GAGGCGCGAGACGACGGTCCAAGGCCTTCGTGGGCTCTTTGCCCCGCAATCCGGAAGCTGCGC
GGCAGTTCGTCGCCCCTGCAATGCTGCCATTGTAGAGCTCCTCCGCTCGCCGCGGCCTCTTTT
CTTGCGGCCCGTCCGCGATTGACCTCACATCCTGATCCCTTCTTGCGTCGTCCAGAAAGTGAT
TGACGGCCAGCGCCGCGCTTGAGATCTTCCGGCGCGCGGCGATTTCATCGCTCCGGCGCGCCG
TGACTGTCACCTGCGAAGGGATTATAATGAAACATAACATTGGGTGGCTTCTACCCGCCGCCC
TCGCGACGCTTGCCTTCGTCCCGGCCTGCAGCCCGAATCACGGTGAGGATGCGCCCTCCGTGA
CGTCAGCAGAGAGCGGCGCGGCGCCGAGCGCTGACTGCGTCGCGCTCGGGGCGAAGCTCCAGG
CGGCGCTGGACGGCGCCGCCGCCGCGCAAAAGGCTCCGGGAGCCGCAGCGGCGGTCCAGAGCG
GGGACTGTGTCTGGCGGGGCGCCACGGGCGTCTCGGACCTGGTCGCGAGCACGCCGACGAAGC
CTGGAGATCTCTTTCGGATCGGCAGCATCACCAAGACCTTCGTCTCTACGCTGATACTCATGC

TCCGGGCAGAAGGCCGGTTGTCGCTCGACGACGCGGTGTCGAAGTATGTGAAGGGCATCCCCG
CCGGCGACCAGATGACGCTGCGCCAGATCCTCGGTCACACGAGCGGGCTCTTCGATTACACGT
ACAGCCCGGCGCTCGGCCAAATGATCGAGGTGGATCCGACCCGCGCCTTCGCGCCGGCAGAGC
TCATCGCCCTCGCCACGGCCGAGGCGCCGTATTTCGCGCCGGGCGCGGGTTTTCGCTATTCGA
ACACCAATTACATCGTGGCCGGCCTGGTGGCCGAGGCGGTGTCGGGCGGGACGCTCGCCGGGC
TGCTCCGCACGCGCATCCTAGACCCTGTGGGCCTCGCGCACACGTATCTGGACGGCGCCGAGC
CGCCGGTCCAAGGGCTCATCCGCGGCTACGGCGACTACGGCGCGGGCTTGGTCGACATCACCG
ACCAGCTGTCGCCCACCGAGGCGTGGGCCGCCGGCGCCCTGGTGTCGAACGTCGATGACCTCA
ATCGCTTCTTTGCCCTGCTCATCAGCCACGAGCTGCTCTCGTCGGACGAGCTTCAGGACATGA
CCACCTGGACCCCGACGATGTGGCCCCACGAGCCCGGATATGGCCTCGGCCTCATCGAGCGCG
ATTCTGCGCTCGGCTCCCTCAACGGGCACTGCGGAATCATCTGGGGCTTTCAATCGGCGTCGT
ACGGGGTGCCCGGCCGCGGCGACGCGATCACCGCGCTCATCAACCGGAGCGACGGCGACGCAG
CGCGGCTCGTCGACGAGCTCGCGAAGGTCGTGAAAGAGCGCTGATCGAGGCGGAATGGGAGCG
CTTCGGCGGGTGGTGATGGCGCCCGGCGCTCAGAACGCGACGCGCAGCCCCGCGCTCAGCGGG
CCTGCGCCGGGCGACGCGGCCACGGCGCCCGGACCGACGAGGAGCCGCGCGACGGCGGGCGCG
CTCGGCGCGTCGTCTCGCCGCACCCGCCGCTTGCCGAACACGTAGAGCGGCAGGCCGACGGCG
ACCCCGGCCACCCCGCCGAGCGCGGTGGCGATCGCCACCTCGGACGCCTCGGCGCGCGCGGCG
CTGCTGTCGTCGTGGCTCGCGAAGACCAGCACCGCGCCGCTGAGGATGGCGGCGCCGCCCAGG
GTCGTGAGGACGAGCCCCGAGATCACCATGACCGGGCTGTTCCACTCCGTCGTCCGCTCCTCG
AAGTCGCGGAACGCCGCCCTCGCCGCCGCGAGCTCCAGCTCGATCCGGCGCTGCTCGGCGCGC
CGCTCGTCCGCGGAGCCGATCTCGTGGACGCGGCGGGCCTGGCCCTCCAGCGCCGCGATGCGC
GCCTCGTGCGCGGCGGCGGTCTCCTCCCACGTGGCCCCTGGCGGGACCCCCGCCACGGCCGGC
GCGACAGAGGGCGCCGACGCCGGGGTCGAGGCGGGCGCCGCGGGCGGCTCCGCGGCCACGGAA
GGCGCCGCCGCCGCGGGAGGCGCGGGCGGCTCCGCGGCCACGGAAGGCGCCGCCGCCGCGGAA
GGCGCGGGCGACTCCGCGGCCACGGAAGGCGCGGCCGCCGCGGGAGGCGCGGGCGGCTCCGCT
GCAACGGAGAGCGCGGGCGCCGCGGCAGCCAGGCCCAGCGCCCACGCGACGACACGACGGCGC
GCCGCAACCGCGCGCGGGCGCGCGAAGCGGAGGTGGACCTGCTCCATGCGCGCAGCGTCGCCC
CTCGACAGGGCCGGGTCAAGGCGCGGGAGTCCGAGAGCACGAGACCTCCGCGCCGCAGGAAAC
AGGCGCGCCGGCGGCCCGCGCGGCGGCTCGCCGCTCACCCCTCGCGCGGCCGGCCGCGGCGCC
GCCTCCCCTCCCCGGCGGGCCGCGCGTCGGCGGCCACGCGGAGCAGCTCCTGGAAGTGCCGCT

CCACCGGGCCGAGGTCGATGCCGTCCATGAACGACGTGAACGCGAAGTACGGCAGCAGCGTCT
GCCAGGCGCGCAGCCAGCCCGGGAGGTAGCGCGGCCGCTCCAGCACGCCGGCGGCGCGCAGCT
CCGCCAGCACCGCGATGTCGTCGAGCGCGATCCGGCCGCACACGAGCAGCTCCACCTCGTCGC
GCAGCCCGCCGCGAAGGACCGCGGCCAGGAAGGCGTAGACCTCGAGCAGCCCGGCCAGGTAGC
AGGCGTCCTTGGTGAACGGCGCGCCGCCCTCGACGAGCCCGCCGCGGCACACGCGCTGGGCGT
CGAAGTAGGCGTCGCGGCGCTCGGCGCCGCGCTCGCGCAGGTGCCGGTACAGGTCGAGGAAGC
TCGCGCCCTGCTCGGCCATGTCCACGAGCCGCACCCGCTCGGCGAGCCGGGTGAGGCGGCCGA
TGGAGAGCGAGCGGCTGTAGAGCTCGGCGAAGATGGCGAGCCCCTCCTGCGTGCGCGTGGTGC
GGGGGCCGCCCGAGCGCAGGAACGCGCACCGCGGCTGCGCCGCGCCGTTGTGCGCGGTGAGCG
CGTGCGTCTCGACCTCGTGGTGCCACAGCCCCTCGGCCTCCCACGCCGCGAAGGTCGCCTCCG
GCCGGATGCGGACCCGGCTCATGCCGGCGACCACCTTGGCCGTGACGCGCGGGTCGACGGTGA
TCTCGAGGTCGAGCCGCGGCGCCCGGCCGGCCACGCGGGCGGCGAGCATGTCCCGGAGCGCGC
CGGCGTCGAGCGGCTCCTCCTCGGGATCGCTGGCCTCGTCCCAGCCGTGGACGCGCAGGCGCT
CGGTGAGGTGCTCGGCGAGGTCGATGTTCCTGAGCGAGCCGCCGAAGAACCGCGAGCGCGCGC
CGCCGTAGAGCTCCTGCGACCGCGCGGAGAACGCGCGGGTGCCCGCGGCCTCGAGCAGCTCCG
CGGCCTGGATCTGCGCGCGGACGTTGTCCCGCAGCCAGCCGAGCGCCGGCGCGTCCCCGTCGA
TGGCCCCGAGGAGCTCGCGCAGCTCGGCGACGCGCCGCGCGAGGCCGTCGCGATCGACGCGGT
ACTCGACCTCGGGGAGGCGGTCCTCGCCGGCGGCGAAGAAGCGCTCCTCCACCTCGCGCGGCC
AGGCGATGTCCTCGAGCAGCTTGAGGGCCTTGCCCTCCGCCAGGCGGCCGCCCACCCGATCGA
GCTGCTCCAGCACGGCGCGGTCGATGCTCATCGAGCGCAGGATCGCCGAAACCGCGAGACGCC
GGAACCGTCATTCCCTCGACGAGGCAGCGATTGCCATGTTCCGTCGCTTTTTGGAGCGCCGTC
GTCGCGCTCGCCTGCGGGCTCCGGCGATCCAGCGCGGTTGCATGCAGCGAGGGTGTTCCGGGG
CTGGCTCGAGAGCGTCCTTTGGCCCACACCCGAGACACGAATGCTCCGCGCCGAGCGCGGTTG
ACCGTGGACCCGCCGGAGAGCCGATGATACGGTCCGGCCGATGTCGGAGAGTGTAGCTCAACT
CGAAGAACACCGCGCGGCGCTCACCGGACACTGCTACCGGATGCTGGGTTCGGTGGTCGACGC
CGACGACGCCGTCCAGGAGACGATGGTGCGCGCCTGGCGGAGCCTGGATAAGTTCGACGGGCG
CTCGTCGCTGCGCACCTGGCTGTACCGCATCGCGACGAACGTCTGCATCGACCTGCGGGCCGA
CCGCGCGCGCCGGGCGCGCCCCATCGAGGAAGGCCCGGTCGGCACGGTGGACGACGCGCTCGA
GACGCGCCCGCGCACCCACTGGCTGGAGCCCGTCCCCGACGCGCACGCCCTGCCGGCGGACAT
CGACGCCGCGGAGCGGGCGATGCTCCGCCAGAGCATCCGCCTCGCGTTCGTCGCGGCGCTCCA

GCACCTGCCGCCGAAGCAGCGCGCCGCGCTGCTGCTCACGGAGGTGCTCGGCTGGTCCGCCGC
GGAGGTCGCCGACAGCCTCAACACCTCGGTCGCCGCGATCAACAGCGCGCTCCAGCGCGCGCG
GGCGACGCTCGCGAGCCGCGATCTCGGCGACGCGCGCCCCTCGCTGCCGGAGCCGCAGTCCGC
GCTGCTCGACCGCTACGTCAACGCCTTCGAGCGGTACGACGTCGACGCGCTCACGGCGCTGCT
GCACCAGGACGCGACCCTGTCGATGCCGCCGTTCACCCTGTGGCTCCGCGGCCACGAGTCGAT
CCGCGCCTGGCTCGTGGGCCCGGGAGCGGGCTGCCGCGGGTCGCGGCTCATCCCGACGGCGGC
GAGCGGCTCGCCCGCGTTCGCGCAGTATCGCCCGGCGCCGGAGGGCGGCCACCGGGCCTGGGC
GCTCATCGTCCTCGACGTCGCGGGGGACCGCATCGTCAGCATGACGTCCTTCCTCGACACCGA
GACGCTCTTCCCGCGGTTCGGCCTGCCGCTCGATCTACCGGCGTAGCCGCGGGCGCCCTGCCT
GCCTCGCCGCGGGTGCCCTGCCTGCCTAGCCGCGGGCGCCCGGCCTGGCCACGGGCGCCCGGC
CTGGCCACGGGCGCCCGGCCAGCGACGGGGCGACGATTTTTTTCTGAGCGACCGATGAGTCCT
GACGGGGCCGGGGGTCTACGGGGGTGAATCCAACACGGAGGCACCCATGACCGTGACCATCGC
CAGCATCGATCATCGTGACCAGGACCTCATGACCGGGCCCCAGGCCAAGGCGCCGGCCCGCGC
GGCGGCGCCCGACGCGGCGCCGTCCAGGCGAGCCGTGTGGGCGGGCCGCGTCCTGAGCGGGCT
GGCCACGCTGTTCCTGACGTTCGACGCCGCGGTGAAGGTGCTGAAGCTGTTCCCCGCGGAGGC
GTCGACCGCCGAGCTCGGGTTCCCGGCGCACCTCGTCCCCACCCTCGGCTACCTCCAGATCGC
TTGCCTCGTGGCCTACCTGATCCCGCGCACCGCGGTGCTCGGCGCGATCCTGTGGACCGGCTA
CCTGGGCGGCGCGATCGCGATCCACGTGCGGGTCGAGAACCCGCTCTTCAGCCACACGCTCTT
CCCCATCTACGTCGCCGCGTTCCTCTGGGCGGGGCTCTGGCTGCGCGACCGCCGCGTGCGCGC
GCTGACCGCGAGCCCGTCGTCGCAGGGCCGATGAGCTTCACGTTTCACGAGAGTCCATCACGG
TAAAAGGAGAAGCGAGCCATGACCACAAAGAACCCCCGCAAGCTCTTCGTCAACCTGTCCGTC
CGCGACCTGAAGCGATCGATGGAGTTCTTCAGCAAGCTCGGGTTCGAGTTCAACCCGCAGTTC
ACGGACGAGAAGGCCGCCTGCATGGTCGTCAGCGAGGAGGCCTATGTCATGCTCCTCGTGGAG
TCGTTCTTCAAGACGTTCATGAAGAAGGAGATCTGCAGCACGAGCACGCACACGGAAGGGCTC
TTCGCGCTCTCGTGCAGCAGCCGGGCCGAGGTCGACGACATGGTGAAGAAGGCGGTCGCGGCG
GGCGGGTCGCACGCGATGGATCCGCAGGATCACGGCTTCATGTACGGGTGGAGCTTCTACGAC
GTGGATGGCCACCACTGGGAGGTCATGTGGATGGATCCCAAGGCGATCCAGCCGTAGCCGACG
GGGCTGGGCGCGCCGCCTGGAAGAGCCCCCGTGAGGCGGGGAGGCGGGAGGATCACCGTCTTC
GTAGCCCACAGCGATGCAGTATCCGTCGCGCTTCGTATCGAAGCACGGCTGTTACGGGCGCGT
CAGAGCGCGTCGCAGGTGATGCCGAGCCGCAGCAGCGACACGGGCACGAGCGTGGCTCCGATG

GAGATGAGCCGAGTCTCGCCCATGGTCTCGGGGTCATGAATGGATGAGTAGGGGACTCGCTCC
TTCGTCACGTCGTGCTCGACGGCGACGGCGAGGCCGAGCTCGAAGTGCACGGGGCCTGGACCG
AAGATCCAGCTCGCCCCGGCGCGAGCCCCGACGAAAAGCGTGTCGCCGTCGACGCCAGGGCCG
TCGTCCCAGCCGGGCGATCCCACCGCGGTGTAGGTGTGTTTCCCGAAGGAACCCGCGAGCGAG
AGTCGAAGTCCGACCGGCGCTCGCCACGCGACGCCCGCTGTCGCGCCGACGCCGCCGAAGCTC
TCCCCGAAAGGCTTATCCCCTGTCTCGATGAAGCCACCCACCTCGATGACGCTGATGCGGTAC
GTGAGCGCGAGATTGAGGTGCACCCCAGCGCTGTCCGAGCCCGAGTAGAGGCCGGCGCCCACC
TGCACGCTGAAATCCATGCTCGGCGCGGATCCGCGCGCAGGAGCGACGCCAGGGGCGCTGCCC
TCCTGCGCGCGGGCCGTCCCGACGCAAAGAAAGAGGGCTGTCGCGAAGAATCCAAGCGAGATC
GATCGAAGTGAGCGCATGTCGGGCCCTGGAGCATCCGCTGTACCAGGTGCGTCGTATTCATGC
GGCGCGCCGCCGGGCGCCGCCGCGCTGGCCTGTCCGACGCGAGATCACGAATCCGCCATCGCT
CCCCTGGGCCGCCGGCCGCTCTGGTTCGCCTGCGGGCGTGCGCCGGCGCTCGTGTGGCCCATG
GCAACCTTGTCGCGGTGTCGCTCGAACAGCACAGAGAGTATCGCGTCCGCAACAACCGCGCGA
CCCGGCGAGACGCTCGTGGGGCCCCCTGCCTCCCCACTTCATCATAACGCCATCAGGAGCACT
CGACATTTCATTTCTTCACCTCCACTGGCTGAGGGCGACGGTGCTCGTCATCGGCCGGTTGCT
CTGGCGGTTGCTCTGGCGGGGTTTCTGACGCCCGGAACTAACGCTTCGAGCGCTCCCCCTTGC
TCTCCCGTTCCTTCAGCTCCTCCAGCAGGTCGTCGAGGCGCTCGTAGCTGCCTTCCCAGAAGC
GGCGGTAGTTGTCGAGCCAGCCGCTGGCGTCCTCGAGCGGCTTGGCCTCGATCCGACAAGGCC
TCCGCTGCGCGTCGCGGCCGCGCGAGATCAGGCCCGCTCGCTCCAGCACCTTGAGGTGCTTGG
AGATCGCGGGCTGGCTCATCGCGAACGGCTTCGCCAGCTCGGTCACCGACGCCTCGCCGGACG
CGAGGCGCGCGAGGATCGCTCGCCGTGTCGGATCGGCGAGCGCAGCGAACGTTGCGTCGAGGC
GCTCGGACGGGGTCATTGCATAACTCCTTGGTATAAAAACCAGTTAGTTATACAACCTGGGGC
CCGGGCGGTCAAGCCTCCAGGCGATGGCGGTTCGGCCCGGGGGCTCCGCTCGCGGCACGCGCG
CCGCGCGGCTACGTGCGCGGCGCGGTGAGCACGTCCTGCAGCGTGGCGCCGACCACGGGCTTG
GTCAGGTGCAGGTCGAAGCCGGCCCGCCTGGACCTGGCCTGATCGTCGGGCCCGCCGTAGCCC
GAGAGCGCCACCAGGTAGAGCGCTTCGCCGCCGGGCGCGGCCCGCGCCCGGCGCGCGACCTCA
TAACCGTCGATGCCGGGCAAGCCGATGTCCACGAAGGCCACCTCGGGGCGCAGCTCCAGAAGC
TTCTTCACGCCCTCCAGCCCGTCCACCGCCACCGTCACCTCGTGCCCCAGCGCCTCGATGTAC
GCCCGCATCACCCGGCGCACGTCCTCCGCGTCCTCCACGACGAGCACCCGGCGCCGGTCAGCC
GCCGCCTCGGGCGCCTCGGCGCGCTGCGCCGGAGGCGGCGGCTCGTCGCGCTGCGCCGGA

GGCGGCCCCTCGCGCGGCGGGGGCGGCCCGGCGCTCGGGGCAGGCTGCGGCGCCGCCCCGGGG
CCGAGCGGCAGGCGCACGGTGAACTCGCTGCCCTGGCCCGGCCCGGCGCTCGCCGCGGCCACG
CTGCCGCCGTGCAGTTCCAGGAGCCGCCGCACCAGCGTGAGCCCGAGCCCCAGCCCGCCCGTG
CTCCGGTCGATGGTCTGGTCGACCTGCGTGAACAGATCGAACACCTTCTCGAGCATCGCCGCC
GGGATGCCGCGGCCCGTGTCGCGCACCCGCAGCACGGCCTCGGGCGCGCCGACCGCCGCCTCG
CGCGTGAGGCGCACCGAGATCGAGCCCCCGGCGGGGTGTACTTCGCGGCGTTGGTCAGGAGG
TTCGTCACCACCTGCTCCAGCCGCGTCGCGTCGGCCCGCATGCCGAAGTCCCCGGGCCCCACC
GACAGCGACACGTCATGGCGCCGGGCCTCGACGGCCGGCCTCACCGCGGCGGCGGCGCTCTGC
ACCACCGCCGCGAGATCGACGTCCTCGAGGCGCAGCTCCACCGTGCCCCGCGTGATGCGCGAC
ACGTCGAGCAGATCGTCGACCAGCCGCACGAGGTGGCCCATCTGCCGCCGCGCGATCTCCCGG
TAGCGCGCCGACGCGGGCCCGTCGCCGTCCGCGTCGTCGAGCAGCGTCAGCGACAGGCTGATC
GAGGCCATCGGGTTCCGGAGCTCGTGCGCGAGCATCGCGAGGAACTCGTCCTTGCGCTGATCG
GCGAGCTTCAGCGCCTCGACGAGCGCCTCCACGCGCCTCCGGGCGCGCACCTGGTCGGTCACG
TCGAACGCGAACACGAAGACGCCCTCGACCGCCCCGTCGCGATCGCGCATCGGCTGGTAGACG
AAGTTGAAGAACACCTCCTCCGTCGTGCCGTCGCCCCGGCGATCGAGCCGCACCGGGAGCTCC
TTGCCGACGATGGGCTCGCCGGTGCGGACCACCGCGTCGAGGAGCTCCCAGATGCCCTGTCCC
TCGAGCTCGGGGAGGGCGGCCCGGATGGGCTCGCCCACGAGCGATCGACCGCCGACGAGCCGC
TGGTAGAGCGGGTTGACCACCTCGAAGACGTGCTCCGGCCCGCGGAGGATGGCGATGGGCCCC
GGGGCCTGCATGAAGAGGTCGTTCAGGTACTGGCGCTGCCCCTCGGCCTCGCGCCGGCGGCGC
GCGAGCTCGACGTGGATGCGGACCCGCGCGAGGAGCTCCTTCGCGGAGAACGGCTTCACGAGG
AAGTCGTCGGCGCCGGCCTCGAGGCTGTCGACGCGCGCCTCCTCGCCCGCGCGCGCGGAGAGC
ATCACCACGGCGACGCCGCGGGTGCGATCGTCGGCGCGCAGCGCCCTGAGCAGGCCGAAGCCG
TCGAGCCGCGGCATCATCACGTCGGTGAGCACGAGATCCGGCGGGTGGGCGCGGGCGCGCTCC
AGGGCGGCCCGACCGTCGGCCACGGCCTCCACCGTCCACCCCTCCGCCACGAGCAGCCGCAGC
GCGTACTCGCGCATGTCCGCGTTGTCGTCGGCGACGAGGACGCGCCCCGGCAGCCTCCCGGCC
GGCCCCTCGCCCGCCGGCCGGGACGCCGGCGCCTGCTCGCCGCGGAGCCACTGCGCGGCCTCG
TCGAGGAAGGGCGCGGCGTCCCGCCCCCCCGCGGCCGGCGCCGAGGCCGGCGCGAC

or its complementary strand,

(b) DNA-sequences which hybridise under stringent conditions to regions of DNA-sequences according to (a) encoding proteins or to fragments of said DNA-sequences,

(c) DNA-sequences which hybridise to the DNA-sequences according to (a) and (b) because of a degeneration of the genetic code,

(d) allele variations and mutants resulting by substitution, insertion or deletion of nucleotides or inversion of nucleotide segments of DNA-sequences according to (a) to (c), wherein the variations and mutants offer isofunctional expression products.

7. DNA sequence according to claims 1 to 5, wherein the DNA is selected from the group consisting of

(a) the following DNA Sequence:

Seq ID No 2 (>pEPOcos6 region)

GGATCACCTGCGGCGCGATCGCCGACCTCGTGCTGGTGTTCGGCTCGCTGGATGAGAAGCCGG
CGGCGCTACTGATAGAGACGGCGACGCCCGGGCTGCGGGTGGAGCGGTTGCGGGAGATGCTCG
GCTTTCGGGCGGCCCACCTGGCGAAGCTGTCCTTCGACGGTTGCGAGGTCCCCGAGGCTCAGC
TGATTGGCCGGCCCGGCTTTGCGCTGATGTATCTGGCCCCCTACGCCCTGGATTTCGGTCGGG
TCAGCGTCGCCTGGGCCTGCCTGGGCATGATCCGCGCTTGCCTGGAGACCTGCGCACAGCACA
TCCTCACCCGCCGCACCTTCGGCCACCTGCTAGCCGATCACGGCATGATCCAAACCCTGATCA
CCAACCTGGGGATTCACCACCAGGCGACGCTGCTCCACACGCTGCAGGCCTGCCGCGCCAGGG
ATCGCGGCGACGTGACCGCCTCCGAGGCCACCCTCGCCGCCAAATACCTCGCGTCGCGGACGG
CGGTCCAGGAGACGACCAACGCGGTCCAGATCATGGGCGCGCTGGGCTGCGACGAGGAGGGCG
CGATCGCCCGCCACTTCCGCGACGCCAAGACGACCGAAATCATCGAAGGCAGCAACCAGATCA
TCGAGGCGCTGCTGGCCAAGAACATCGCCCGCGCCGGTCGCGACAACTATCGCCGCTTCCTCG

ATGCGGAAGTCGAGCCCGGTCGGGCCGGAGGCGCACCATGACGAGCGCGGTCCCGACGCGTCA
AACCAGCCTGCTCGACGACTTCGAGCGCGTCGCCGACGTCGATCCAGAGCGGATCGCCGTCCA
CGCGAGCGAGACGAGCCTGCGCTATGGCGACATGAATGCGCGCGCCAACCGCATTGCCCACGG
GCTACGGGCGCGCGGGATCGGGCCCAATCAAATCGTGGCGGTGGCGATGGCCCGCACGCCCGA
GCTGATGATCGTGCTGTACGGCATCCTCAAGGCCGGCGCGGCCTACATGCCCATCGCCCGCGA
CGCGCCGCCGCTGCGCCGCGATCATATGCTGCGCGAGAGCCAGGCTGCTCTGATGATCGCCGA
CGAAGAGATCGCGGGACTCGCGGCCCGGGTGCTGACGCCGGCCGACCCGTTCTTCGCGGCCAT
GCCGGACCACAACCCCGAGCCGCGTCACGACCCGACCGACCTGATTTACGTCATCTACACCTC
GGGCTCGACCGGCCAGCCCAAGGGCGTGGCCATGGAGCACCGCGCCGTGTGGAATCGCCTGAC
TTGGATGCAGGCCCAGTATCCAATCGACACGCAGGACGTGATCCTCCAAAAGACGCCGATCGT
CTTCGACGTGTCGGTCTGGGAGCTGTTCTGGTGGCCGCTGGCCGGCGCCTCGGTGGCCCTGCT
GCCGCAATCCATGGAGAAGTTCCCCTGGGCGATATCGGCGACGGTGGCGCGGTGCGGGGTGAC
GGTGATGCATTTCGTACCATCGATGCTGATGGCCTTCCTTCAGGTGGTGGCGGGCCGGCCCGA
GATGGCGGACCAGATGAAGGGCCTGCGCTACGTCTTCTGCAGCGGCGAGGCCCTGGCGCCGGC
CCACGTGTCAGCCTTTCAGGAGCACATCAACCGAGCGGGCAGCATCAGCTTGACCAACCTCTA
TGGACCCACCGAGGCGGCGGTCGACGTCAGCTACTTCGACTGCCCGCCCGGCGCGTCACTCGC
GCGGGTGCCGATCGGACGAGCGATCACCGGCATCCAGCTGCTGGTCATGCGCGACGGCGTGCC
TCAGCCGCCCGGCGTCGAGGGTGAGCTCGCCATCGGCGGCGTTGGTTTGGCGCGCGGCTACAT
CTCACGGCCAGACCTGACCGCCGACCGGTTCGTGCCGCATCCAGGCGGCGACGGCCAGCGGCT
CTACCGCACCGGCGATCTGGTGCGCAGGGACGCGGACGGCGAGCTGGTCTTCCTGGGGCGCAT
CGACCATCAGGTGAAAATTCGCGGTCTGCGCATCGAGCCCGGGGAAATCGAGGCCCAGATCAG
CGCCCATCCCGATGTGGCCGACTGCGCGCTGATTATCGAGCAGGACTCGGAAACCCTGCCCAA
GCTGACCGCCTACATTGTCGTGGCGCGACCGGGCTTGACCCGGAAGGCGCTGCTACAGTTCCT
GGGCGCGCGGCTGCCCGACTACATGCTCCCGAACCGCTTCCTGACCCTCACGGAGCTGCCCGT
GACCGCCAACGGTAAGCGCGACTGGCGCGCGCTGCTCGGCCCGCTCGAGACCCTGCCTCTCCC
TTTCTCCTGAATCCAACCAATACGAGGGATTCATGTTACACCCGATTCCCACCGACCGTTTCG
CCCTGAGCCGACCGCTCTTTCGCGGGTACCTCGCGCACGATCCGATCGTGCAGGCGTGCTGG
CGGGCGACCATCCAGGCTGGGTCCTGGTGGACCGCGAGCCCGAGCCGCGCACGGCGCTGCTGT
GGGCCTTTTCCGATCGGCTCTTCTGCGTGGGCGCAGCTGACACGCTGACCCCGCACGCGCTGG
CCGAGCTGTTCCACGACCGACTGATCCCCCAGGCCCGTAAGATCGGGCAGCCGTTTTTCCAGG

TTCAGGGCGAGACGGTCGACACCTGGTCGGACCACCTGCATCAGGTGTCGCCGCACGCGACAG
TCTCCTTCCGCCAGGCATTCCGCTTCGACCGCGACCTCTTCGAGCGGCTGCCAACCAAGCCGG
AGCTGGCAGAGGCGCGGCTCGTGCCAATCGACGCGCGGCTGCTGGCCGAACAGGCTGATCTGC
GCGAGCGGATACTGGCCTCCTGGTCCAGCGAAGCTGCCTTCCATGCGCGCGGTTTCGGCTTCT
GCTACCGCGTAGGTGACCAGCTGCCGAGCGTGTGCCTGGCATCGCACGTAGGCGGCGGCGCGG
CCGAGCTGAGCATCAACACCGAGCTCGAAGCGCGCAATCGAGGTATGGCAACGCGGCTGTGCC
GGCGTTTCATCGCCGAATCGCTGCAGCGCGGCCTGACGCCTTGCTGGGGCACCGAGACCTTTC
GCCTGCCGTCAATCGCGCTGGCCCAGAAGCTCGGTTTCATCCCGACCTTCACCTTCCCCACCT
ACTGCTTCGCGACCGGCACCGAACAGCCGGACGACAACTTCCTAGGCGAGCTGTACTACAGGG
AATCGCGCATCGCCGGAAGTGGGACCGATGAGCCGCAAGCGGTTCGGCTGGCGCGGGGTTGGA
GCCTGGCCGGCGACACCGAGCGTGCCGCGAGCTTCGCCGCACGCGCCCTGGCCGAAGGGTGGG
CCGGCCACTCGACTCTGGCCACCGATCCGGATTTCGCCCGATTGCGCGCCAGCGCCGCCTGGC
CCCGCCTCAATGTCCCTTGAAAGGTCACGTGGACTCATGATGTCCCCTTGAAAGGTCACACTC
CGAGTCATGATGATTTGTCACTCCCACCGCTTCATTTTCCTCCACGTTCCCAAGGTCGCCGGC
ACAAGCGTCAAGGACGTCCTCGGCCAAGAGCTATTCCAGGAGGACCAGGTCACGTTCCAGATC
GCTCCCAATCCCCACTACCCACCTGAATGGACTGCGCCTTACGAGGAGCACATTATTGCCGCT
GAATTGAAGAGCCAGTTGGCGCCGGAAATTTGGGACGATTACTTCAAGTTCGCCTTCGTGCGC
CATCCGCTCGACTGGGCGGTCTCCAATTACTTCTTCTTCCTGCGCGACCGCAAAGGCCATCCG
GCCCACGAATTCCTGGAGCGGAAGGGCTTCGCCGGTACCATGGACATGTTTTTCGGAGCGGCC
GGGCGCCATCCGCTGGTCGCCGGCATGCGCTTCAGCCAATGGGAGTTCTTGTGCGACAGCGAG
GGCCGGACGCTGGTGGACTTCGTTGGCAAGTACGAGCGGCTCGAGCAGGACTTCGCCGCCGTG
TGTATCCGCATCGGGCTGACCCCGCCCGACTTGCCGTGCCTCAACCAGACTCGCCACCAATCC
TTTACCAGTTACTACGACGAGGCTTTGATGCGCCAAGTCAGCCGCGCGTTAGCTCGCGATTTC
GAAATTTTTGATTATGCCTGAGGCGGACCCGTTGCTTCGCCACCGGTGGATTATTCGATAAGT
TATTATATTTTCAGTTGATCATGTGAATGTCGATCCAGCCAACGAGGAGGATACCTCCGCGTG
CGGCTATGGGGGCGCAGAGGTCACGACTACGTGTAGAAATTTGTCGAACACACCACTAGCTGC
CACCGATTGGGAGCTTTGACTTGAAGATGAAAGTGGACAAGCGGAATGTCGACGACATTCTCG
GACTCACTCCGACACAGACAGGCATCTTGTACCACTACCTGCTGGACCCGCAGGCCGACGCCT
ATTTCGAACAATTGACGCTGCACCTGGAGGGGCCGCTCGACGTAGCGCGCTTCCGCCGCGCCT
GGGAGCGCGTGGTGGCGGCTCACGACCAGCTGCGCGCCGTGTTTCGCTGGCAAGGGATCGAAC

ACCCGGTGCAGATCATCCTCAAGCAGCACGTGCCGGACCTGGAGTTGGCGGAGGTCCCGCGCG
ACGCCGATCCGGCAGCCTTCCTGGCGCAATGGGTCGCGGCCGACCGGGCGCGCAAGTTCGACT
TCGAGACGGTGCCCTTTCGCATCGGCCTCTGCCGGACTGATACCCAACATCACGTGATGCTGC
TCAGCAATCACCATATCCTGATGGACGGTTGGAGTACGGGCCTGATTCTGCGGGACTTCCTCG
CCTGCTACGGCGACTCCGAAAACTGGCGGCCACGCACCCGAACGCACTTCAAGGCGTTCATCA
AGTGGCACCAGAACCGGCCACGCCGGGGCGAGGAGCGATTTTGGCGCGACCTGTTGCGCGATG
CGCCCGACGGCGGCTTTCCCCGCCTGGGCGTCGAAGAAGGCACCCGCCACTCGCTTGACTTCG
GCGCCCGCAGCCGCGCTCTCGACGACCGCTTGACCCAAGGCTTGCGCGACATGGCTCGCGACC
TCGACGTCACCCTCGCCGCGATGCTCCATACCGCTTGGGGCCTTCTACTCCAGCGCTACCAGA
ACAGCTGCGAAGTGATATTCGGGACCACCGTTTCCGGCCGCAACGTCGAGCTCGCCGGCCTCG
ACGAGGTGGTCGGCTTGTTCATCAACACGATTCCGTTCCGCTTCTCGGCCGCGGCCGCGACGA
CGCCCGTCGAGGCCTTCCGTGCGGTACAGCGCAATCTGCTGGCGAGAAGCGAGTTCGAAGCCA
CCCCGCTGGTGGACATCAAGGGCTGGAGTGGTCTCGGTCCGGGCGCGGAACTGTTCGACACCA
TCCTGGTCATCGAGAACTATCCCTTGGACCGCGCTATCTTCGAGAGTGATTCCAGCCTGCGGT
TGACCGACCACCAAATCTTCGAGCGCACCAATTACGGGCTGACCCTGACCATCGAGACCTTCA
GCCGGTTGCACGTGACGCTAGCCCATCGCCGTGACCTGCTGGGCGACGCGGCCGCTGAGCGAA
TGCTAGATCATTTCACCGGCCTGCTCCAAGCCATGCTGCGCTTCCCTCACCAGCCGTTCGCGC
GCCTCGAGATGAAAAGCGAACACGAGGCCCACCGCGTCCTGCACCAACTCAACCAAACGCGTC
AGCCGCTGCCGTCCCAATCGGCTTTCCACCAGTTGTTCTTCGAGCAGGCCCAGGCCGATGGGG
CACGACCGGCGCTGTGGTGCGGCGCCACGCGCTGGACCTACGGCCAGCTGCTGGAACGTGCCC
TGCGTCTGGCGGGACGGCTGCAGGAAGCCGGCTTCGCCCGAGGCGATGTCGCCGCCGTCAGCC
TCGGCCCGGTTCCGGATCTGATTCCCGGTTTGCTGGGCCCGCTGTTCGCCGGCGGCGCCTACC
TGCCGCTCGATCCCACCCTGCCGGCCCAGCGCTCGCGGTTCATCCTCGACGATGCCGGTTGCC
GCTTCCTGATCAGCGACGCGCCACTCGCGGGGCCCACGCCGATCCATCCGGACCCTGCCGGCG
CCAGCCCCGTTGACGTCATTTTTGCCTGTCAGGACGGCGCCGCGCAGCCCGCCTACCTGATCT
ACACCTCGGGCTCCACCGGCCAGCCCAAAGGCGTCTGGGTTAGCCACCGCAACCTGATCAACT
TCCTGACGGGCATGAGCGCAATCCTGCCGGTCGCGGCCGACGACGTGTTCCTCTCGCTGACTA
CCGTGTCGTTCGACATTTTCGGGCTCGAGACGTGGTTCCCGCTCAGCCGCGGCTGCACGATCG
TCTTGGGCACGCGCGCCGAGCAGTTGGACCCGGCCGCGGCTGCCAAGGCCATCTCCTGCCATG
GCGTCACGGTTTACCAGGCGACGCCATCGCGACTCCAACTTCAACTGGAGCACCCCACATTTG

TCCGCGCCATCGGCTCCCTGACGACCCTGCTGGTAGGCGGCGAACCCCTCCCAGCCGAGCTGC
TGCGGCGCGTACGCGAAGTGACCGATGCGCGTATCTTCAACCTCTACGGTCCCACCGAAACCA
CCATCTGGTCCACAGCCGGGGAGGTCACCGCGGCGGACGTCCCGGATATCGGCCGCCCGATCG
CAAATACCGGCGTTTTCCTTCTGGCGCGAGACGGCTCGATCCAGCCGCCGGGCCTGGTGGGCG
AGTTGTGCATCGCCGGCGAGGGCGTGGCGTTGGGCTACCACCGACGGCCGGACCTGAACCGAG
AACGGTTTCGCGAGATTCCGCCGGGCCGCCTGCCCTTTGCCGGCAAGCTCTACCACACCGGCG
ACCTGGCCCGCTGGACCGAAGACGGACGGCTCCTCTGCCTGGGCCGTCTGGACGACCAGCTCA
AAGTGCGCGGCCATCGCGTCGAGCCGGGCGAGATCGAGGCAGTGATGGCGCGCCACCCGGCGG
TCACGCAGGCGGTGGTCGTCACGCGGCCGCGCAACGGCGAGCCGGTCTTGGTCGGGTTCTGGA
CTGCGGAAGGTGAGCCGATGCCAGAGGAAGCGCTGAGCGCTTACCTGGCCGACCGACTGCCGA
GCTACATGGTACCCGAACGGTGCATCCTCATGAAGGCCATGCCGCTAACCGGCAACGGCAAGA
TCGACCGGCGCGCCCTACCCAATCCCTTCGCCTTGACCGAGTCGACCCGGCAGGCGGCGCCGC
GCACCTTGGCCCGCACCGCCGGCGAGCATCGGGTTGCCGAGCTGTGGCAGGCCTTGTTGCGAC
GCGAGGCGATCGGCTTGGACGAACCCTTTTTCAGGCCGGCGGGAACTCATTCGGCTTGATTC
GGCTTCACGCCAAGCTGGAATCCGCCTTCGGGAAGTCGTTCCCGATCACCGATTTGTTCCAGC
ATACCAGTATTCGCAGCCAGGCAGAAATGCTGAGCGGCTCGTCCGTCGAGGCGCCGCTCGCGG
GAGCCGTGCCGCAACCCCCGGCCGCCGCCGCCCAAGTTGCCTCCTCGGCAGCTAAATCCCCAG
GGGAGCGCGGCGCGGCAGCGACGTCGAGCGGCCTGACCGCGCAACCGCCCCAACCCCACTTCC
GGCCCATCGCCGTTATCGGCCTCGCCGGCCGATTCCCCGCCGCACCCGACCTCGACGCCTTCC
TTGAACTGCTCACGGAGGGTCGCTGCGGCATTCGCTTCTTCAGCCAAGCCGAGCTGCGCGACG
AGGGTCTCGACGCGAATCGAATCGCGTGTCATAACTATGTCCCGGCCAAAGGTTTCCTCGACC
GGGCCGACCACTTTGATGCCGACTTCTTCGGCATCCCGCCGCGCGACGCAGAAATCACCGATC
CGCAAATTCGGCTTCTGCTTGAGTGCTGCTGGAACGCGCTGGAGCATGCCGGCTACCCGCCCG
GCGGCGGCGAGATCGGGCTCTTCGCCGGCTCCTCGGCCAACTATCACTGGCTCGAATACGTGG
GCATTTCCGAGGAGAGCAGCAATCGATTCGCCGTCATGATTCAAAACGAAAAGGACTACCTGG
CCACGCGGATCGCCTACCAGCTCGATTTGAAGGGCATTGCCGTCACCGTGCAAACGGCCTGCT
CGTCGTCGCTGACCGCGGTCGAGCTGGCCTGCGATGCGTTACACGCCGGCCGCGTGACCATGG
CTTTGGCTGGTGGCGTTGGTCTGACCTATCCGTTGCGCGCCGGATACCTGCACGAGGATGGAA
TGATCTTCTCCCCCGACGGTCGGTGCCGGGCCTTCGACGCCCAGGCGGCCGGCACGGTCTGCG
GCAACGGTCTGGGCATGGTGGTGCTGAAACAGCTCGACGCGGCGCTGGCCGACGGCGATGCCA

```
TCCACGCTGTGATTAAGGGCATCGCGGCCAACAACGACGGCGCGGCCAAGATCGGCTACACGG
CGCCCTCGCAGAACGGTCAGGCGCGGGTGATCCGCGCCGCCCATAGGCTCGCCCAAGTCGCGC
CGGAGACCATCGGCTATGTAGAAGCCCACGGTTCGGGCACGCCGCTGGGCGATCCGATCGAGG
TGGCGGGCCTGACCGAGGCCTTTGACAGCCCGCGTCGCGGCTTCTGCGCCTTGGGTTCGGTCA
AGTCGAATGTGGGTCATTTGGATGCGGCAGCGGGCATCGCGGGTTTCATCAAGGCGGTGCTCT
CGCTGTCCCATCGGACCCTGTTCGCCAGCCTCCACGTCGACACGCCCAACCCGCAGATCCCGT
TCGCCGACGGTCCGTTCCAGGTCAACACGGAGACCCGGCCCTGGCCAGCTGCCGACCATCCCC
GCCGCGCCGGCGTCAGCTCCTTCGGCATCGGCGGCACCAACGTGCACGCCGTCCTGGAAGAGG
CGCCGCAGTTGGCCGAGCACGCGGGGCGGCGGCGCGAGCGGCAGCTGTTCCTGGTCTCGGCGC
GGACTGCAGCCGATCTGGAGCGACGCACCGCGGCGCTGGTCCGCCACCTGGCCGCGCATCCGG
ACCTCGCACCAGATGACGTTGCCTTTACCTTGCACGCGGGCCGCAAACCGATGACCCACCGTC
GTTTCCTGGTCGCCGCCGACCTCGCGGAAGCCGCCGCGCGTCTGGCCGAGCCCGATCCAGTCA
AATCCGCCGCGGCGCGCGCCGACCGCTGCCAGGTCTGGATGTTCGCCGGTCTCGGCTCTCAAT
ACCCCGGCATGTGTGGCGGCCTCTATCGCACCGAGCCGGCCTTTCGCGAGCAAGTCGACCGCT
GTTTCGACCTCCTCGCGCCGCGTTGCGATTTGAAGCCCTCGCTCTTCCCCGAGCCCGATCAGG
CCATCGACGCATCAGCCCTCGCGGCCATCGACACCGCCCAGATCGCCGTCTTCGTCTGCGAAT
ACGCGCTCGCACGGATGCTGGAAGGCTGGGGGCTGCGTCCGGATCGGCTGATCGGTTACAGTT
TCGGCGAATACGTGGCCGCCTGCCTGGCCGGCGTCTTCTCCCTGCCCGACGCCTTGGCAATCG
TCCGCGAGCGTGGCCGGATCCTGGCGGCGGCCGAGCCGGGCGCGATGGTCAGCGTGCCCCTTC
CGGCCGAGCGCGTCGCGTCGCTGCTGGAGCCGCCGCTTGCCTTGGCCATTGACAACGGCCCCT
CATGCGTGGTGTCCGGGCCGGTCGAACCGGTGCGCACCTTCACCGCTCGCATGAAGCGGGACC
GGGTCTGGGTGACGCCGCTCCAGGCCGAGCGCCCGATGCATTCGCCGCTGATGGCCGAGGCCG
GCGGCTCACTGCGCGCCATGTTGGCCGGGTTCCGCCTGAATGCGCCGCGAATCCCGATCTTAA
GCAATGTTACAGGAACCTACCTAACCGACGAGCAGGCCCGAGACCCCGATTACTGGGCCCGTC
ACCTGTGCGGCAACGTTCGCTTCGCCGACGGTGTGCGAACCTTGTTGGCCGAGCGCGATCCGG
TGTTCCTTGAATTCGGGCCGGGCCGCGATCTGAGCTCCTTGGTGCGCCACCAGATGCCGGAAG
GCGCCGACGAGCCGATCGCACTGATCCGTCATCGCGAAGATCCGGTGCGCGACGAAGACCTCC
TGCTCGATGGCTTGGGCCGCTGCTTCCTGCGTGGGGCGACCCTCCACGGGCAGGCCTTGTACG
CCGGCCGAGGCTGCCGCCGCGTGCCGCTGCCCGGTTACCCGTTCCAGGGTCCACGCTGCATGC
CGGCCCGCGCCGGACTGCCCGGCCTGGCGCGACCGACCGTGGGAGCGACCACCATCAGCTACC
```

GACCAGCCTGGAAGCGGGCGCCGCGCTTGGCGGCTGTCGAATCGCTCGCGCCGCAATCCTGGT
TGGTATTCAGCGACGGCAGCGAATTGGCGGGCGAGCTGGTGGCCGGCCTGCGCGCTTCCGGTT
GCGCGACCACCCTCGTCGAAGGTGGGCTGGCGTTCGCGCGCTTCGCGGGCGGCTTCCGCGCGA
ATCCCCGCGAGGAACAAGATCTCGCACAGCTGTTCGCGACCCTGTCGGCCGAAGCGATGCTGC
CCACCCACATCCTGCACCTGCTCAGCCTGCCGTCGCCGGAGCGCGACTCGCCGCTGGCGCGCC
TGGAGCACCTCACCGAGCTGGGCTTCCACCATCTGCTGGCCCTGGCCCGCCAACTGGAGGCGG
TCGGCGCCCCCGAGGTCCGCCTCGCCGTGGTGACAACCGGCCTGGCGGCGATTGGCGGCGAGT
CCGAGCTGCGGCCCGAGGTCGGGCTGTTGCGGGGACCTGTCCGCGTGATTCCCTTTGAATTCC
CGAACTTGCGGCTGCGCCTGATCGACCTCGACTCGGCCGATCCCATCTGGCGTAGCGGTTGTG
AGCCGTTGCTGCGCGAAATGGGCGCTGCCCCGGGACCTGAAGAAATCGCGCTGCGCGGCACCA
GCCGTTGGGAGTTGGGCTACGAGCCGGTCGAGGGGGGCACCGTGAGCACCATCTCCTCGCGAC
TGCGCGAGGGCGGCGTCTATCTGATCACCGGTGGCCTCGGCGGCCTGGGTCTGGCCTTGGCCC
GTCACCTCGCCCGGAAGTACCGCGCCACCCTGATCCTCGCTGGCCGGCGAGGCGCGCCGGCGC
GCGAGCTCTGGCACCAGGCGCCAGCGGAGTTCGTACCGGTCGCAGCTGCGATCGCACAGATGG
AGGAGTGTGGCGCCCGCGTGATTCCCGTCGCGCTCGACGTCACCGACGCCGACCAAGTGAACG
CGTTGTTCGCCACCATAGAAGCTACGGTCGGCAAGATTGAAGGCGTTTTCCACATGGCTGGCA
TCGTTGACGGCGGCATCATTCGAACGCGCACGCGCGCTGCCAGCGACGCCGTGCTGGCGCCCA
AAACGGTCGGAACCTGGATTCTCGATCGGGCTCTCCGCGGCGCCGGTGGCCGCTTCCTGGTGC
TGTACTCCTCGATCAACGCGGTCGTCGCGCCCTTCGGCCAGGTTGCCTACGCCGCCGCCAACG
CCTTCCTCGACGCCTTCGCCAGCGCCCACGAACACGACGAGCGTCTTTTCCGCGTCAGCATCG
GTTGGGACACCTGGCGCGAGGCCGGCATGGCCGTCGATGCCGCCCGCGCCCGCGGCGACCAGG
CCCCGCTCGAAGGGCTTAGCGACGAGCAGGGCTTGCGCCTGCTCGAAAGCGCCTTGGTCGGTT
GCGAACCGCGACTCCTCGTCTCCATCAGCGAACTGCGCGCTCGACTAGCCGAGCATCATCGCA
ACGGCGGCATTCCCCGGTTGCTCGGGCCCCGCGCCAACGAGGCGGGTGCAGCTGATTCCGGCG
AGGAGGGCGCCACGCAAGACGCGTCGCCGGCCCGTCGCGCCCGTCCCGATCTGGTCGTGGCCT
TCGCGCCGGCCGGCAACGAGCTGGAGCGCCGGATCGTGGCCATCATCGGCGCCTACCTGCGGC
TCGGTCAGGTGGGCGTCGACGACAACTTCAACGATTTGGGCGCCACCTCGCTCGACCTCATCC
AGATCGCCCAACGCCTCGGTCGCGAGTTGGGCCGCGATGTCCCTGTCGTCTCGCTCTACCAAC
ACCGCACCGTACGCGGGCTGAGCCGCTTCCTCGGCGGCGCGCTCCAATCCGCGCGGTCCGGCG
TCCCGACGGGCGCTGCCGCACCGGGCGCCGCCACGCCGGGGGTTGCCACCCCGCCGCGGCCAC

AACCGTCGCGCCAGCACCTGGAAAAACGCCGTCAATTGAGGAAAAAAGGGGGGCCTTCCCATC
ATGAGTGAAGTATCCATTCGCCCCGGCTTGGACATCGCGGTCATCGGCATGGCCTGCCGCTTT
CCCGGTGCCCGCAACCTCGCCGAGTATTGGGCCAACCTGATCGAAGGCCTCGAAACGCTCAGC
TTCTTCAGCGAAGAGGAGCTGCGTGAGGCCGGCTGCGATCCGGTCCAACTGGCCCAGCACAAC
TACGTGCGCACCAAGGGCCTGCTCCCTGACGCAGACCGTTTCGACGCCGATTTTTTTGGTTAT
TCCCCGCGCGAAGCCCAGGTGATGGACCCCCAGATCCGCGTCTTCCACGAGGTCTGTTGGCAG
GCGCTGGAGCACGCGGGCTACAACCCGCATCGCCACACCGGCACGATCGGCCTGTTCGCCGGC
GCCGCGCCCAACGTTTTTTGGGAGTTTCTCTCCTATCGGTCCGATGCCGCCAATTTAGGCAAC
TTCACGCTGGGCCTGCACAACAACAAGGACTACCTGAGCTCGCGCATCGCCTACAACTTCAAC
CTGACAGGGCCCAGCTACACCCTGTTCACCGCCTGCTCGACCTCGATGGTCGCCATCCACCAG
GCCGTCCAGGCGCTGCTCAACGGCGAATGCGACCTGTGCATGGCCGGCTCGGTCTCCATTACG
CTGCCACTGGTTGCCGGCTACACCTACACGCCGGGCATGATCGTCTCGCCCGACGGCCATTGC
CGCACCTTCGACGCAGGCGCCAATGGCACTGTCTACGGCGACGGGGCCGGCGTGGTCGTTCTC
AAGCGGGCCGAGGATGCGTTGGCCGACGGCGACCACATATTTGCGCTCATCAAGGGCTCGGCG
CTCAACAACGATGGCAGTCGCAAGACCGGCTACACCGCGCCCAGCGTGCAGGGGCAGGTGGAG
GTGATCCGCGCGGCGATGAACCTGGCGGAGGTCGAGCCGGAGGCGATCAGCTACGTGGAAACC
CACGGGACGGGCACCACGGTGGGCGATCCGCTGGAGTTCGAGGCGCTAAAGGAGGCCTTCGGA
GGTGGCTGCAAGGCCTTCTGTGGATTGGGTTCGGTCAAGCCGAACATCGGCCATCTGGACGTG
ACGTCGGGGATCGCGAGCTTCATCAAGCTGGTCCTGGCGCTGGAGCACCGCATCCTACCGCCC
ACGCTCCACTTCCAACTGCCCAACCCGAAGATGGATGTGGTCGATAGCCCCTTCTACATCGTG
GCTGAGCGCGAACCCTGGCGCGAAGATCTGCTGCCGCGTCGGGCCGGTGTCAGCGCGTTCGGT
CTGGGTGGCACCAACGTCCACATGATTTTGGAGGAGTTTCAGCGCGAACCGGCGGCGAACAGC
GCGCGCACGCGCCACCTGACGGTGCTGACGGCGCGGTCGCCGCAAGCCCTGGCGCAGCTGGCG
GCCAACCTCGCCGAACACCTGCGCGAACACCCCGAGTTGGCGCTGGCCGATGTGGCCCATACG
CTGCTGCACGGCCGCAAGCCACATCCATTCGCGCGCATCCTGGTGGCGACCGATACGACGGCG
GCGATCGACGCCTTGATGAACGACCGCGATCCGCGAACGCGTTTCTTCGAAGCGACCGGGCGC
GGCGAGTCGGTGATCCTGTGTTTTGACGAAACGCCGCCGGAGCCGCGAAGCGCCCGCTACCTC
TGGGATCACGAGCCGCTTTATCGCGCGGCGGCGACGTCGTGCTTGGCTGGTGAGGTCGCCGAC
CCGGATCTGGAAGGCTGCTTTACTGCCCTGATCGCCGAGCAGGGCGCGGCAGCCGCCTTTTGC
CACCAATACGCGCTGGCCGGATGGCTGCTGGCCATGGGGTTGACCCCGTCGGCGTTGATCGGC

GTGGGCCAGGGCGAGTGGGTAGCAGCGGCGCTCGCGGAGGTGTTCCCGCCATCGGCCTGCTTG
CGCTGGATTAGGTTCGGCGAACGGCTCCCGCAGCCGCGCGATCAACGGATTCCGTTTCTCTCC
AATTTCTCTGGAAACTGGATCGTTGGGCGTGAGTTGGCCGACCCGGATTACCCCAGAAAGCAG
AAGGGTAAGCGCTGCATGAAGCGCCGTCGGTCCCAACCTCGGTCAGCTGGTGCAGGATGGGGG
CGATGGAACCGGCTCGGTCAGCTCGTCGCGCGCTGCTCTTCCGCGGAAGCGGAGGCGGGACG
GTGATCGGCCCGAGGGCGAGGTTCATCTCGTCGTCGACGAGCCGGGCGCGGGTGCGCGCCCAG
TACCTGGGGGCGAGCTCGAGGTAGCGGTCCCGCGGCCAGTAGGGCATCGCGCGAATGACGTCG
GCCAGGTAGGCCTCCGGGTCGAGCCCGTGCAGCTTGCAGCTCGCCACGAGCGAGAAGAGGTTG
GCCGCGGCGGAGGCGTGGTCGTCGCTGCCGAAGAAGAGCCAGGACTTTCTCGCAACCGCAATG
GATCGCAGCGCTCGCTCGCTGGCGTTGTTCTCCAGGCGCAGCCGACCGTCGTCGAGGAAGCGC
CGCAACGGCTGCTCTTGGTTGAGGGCGTAGCCGAGCGCGGTGGAGACCAGGCCGCGCTCGCGG
GGACGAGCGTGCTCGGCCCTGGCCCAGGCAAAGAACGCGTCGACCAGAGGGCGGACGACGACA
TCGCGACGCACCTTGCGCTGCGCGGGCGGCAGGTCCGCCAGCGCGCGATCGGCGGCAAAGAGG
GCGTTGATGCGCCGCAGCCCCTCGACACCGAGCTCGTGCTTGCAGACCGCCGCCTCCCAGAAG
TTGGTACGGCAATGCGACCAGCATCCGACTTCGGTCGGGGGCGGACCGCGCTTCTCGTCGGCA
GCAGCGCCTCTTGGTGGTGTGCCGCGGAAGAGGGCGTCATAGATGGCGTGAGCGTCAGCTTGA
ATATACCGAGAGAAGCCGCGGAACATCTCGCAGACCGCGGCGCTGGTATGCTTGGGCTGGTAC
TCGAAGAAGACGTGATCCTTGTCCGCGAGGACGACGAAGAAGTGTCCCTTGCGGCACGGCCCG
GGCTTCTTGTCCTTGCGCTCCTGGATGGGCCCAGGCTGGACGGAGACCCCGGTGGCGTCCGTG
GACAGGCAGAAGGCGGTCTCGAAGGCCTCTTTGCGCGCGGCCTCGACGATGGCGCCCAGGGTC
GCACCGACGTCTTCGGCGTAGCGGCACATCGTGCCGCGATCGAGCGACGCGCCCTGAAGCTCC
AGCTGCTGCTCCAGTCGATAGAACGGGACGCCGAGCAGGTACTTGCTGGTGAGGATGTGCGCA
ATCATCGACGGCGCGAGGAACGACCGCCGGAACAACTCCTTCGGAAGCGGCGTCGTGATGAAG
ACCGTGCAGGTCTCGCCCTTCGGCGCCGGCGGCGGCGCGTCGAGCGAGGGCGCGTCGAGCGCT
GTGGAGGAAGCGCTCGGCTCGCCGGCCGCTGCCGTGTCCTCCGGGCTGACGCTCGCAGCGGGC
GTCGGCGTCGGGGCTTCTCTCGCGACGACCTGGAGCGGGGCCGCTTCCTCCTCGCCCGAACTG
CTCGCATCCGTGACGGACCGCTCGGCCTTGTACACGACGCGTGCGAGCACGATGCGGCGCATT
CCGCCGCGCTCGTAGCCGAGTCGCGAGGTCTCCTCGACCCCGATGCGCGTCGCCGTCGCATCG
AGCTCGGGGCAGGAGAGCTCGATGCGGACGACGGGCAGGTCGGACTCGGACAGGTCGCGACGG
CCCTTGCCGCCGGACCTTCGTTTCGGCCCCTTGGGGTCGTCGTGCTGCCGCTCGTCGCCTGTA

TTGCGCTCGGCGGCGTCGAGTGCCTTCGCGAGGCGCTGGACCTCGAGGAACATCGAGTCGAAC
GCCAGCTGCTCCGCGCTCACCTCGGCGCGCTCCGCCTTGGCCACGAACAGTCGACGTCGCAGA
AGCTGCAGCTGCTCGAGCGCACGGGTGTAGGCGCGCCGAAGCTGCGCGAGCGCATCGCGCGCT
CCCACGAGCTCGCTCTTTGCCGCGGCGAGCTCCGCTTCGAGCTGCGCGATGCGCTGCTGCTCG
GCCGAGAGCGTCGGCTTGGCGGCGGCGTCGTGCACGACGCCGCTCTACGTAAGCCGCGCGTAC
TTGTCGAGCGAATTCGTGCGGCTCAGTGGACGCGGCGCGGTGCGCGCCTTCGCGGTTTGGACG
TGGGCGCGATCTCGATGCCGTCGAGCAGCGTCTCGAGCGTGGCGTCGTCCACCTCGACGTGCG
TGGCGCCCTCGGTCGGGGGGTCGGGAAGTGCGAACGCTCCGCGATCAAGGCGTTTTGAAAACA
GGCAGATTCCACTGCCATCGAAGAAGAGAATCTTGATCGTGGTCCGCCGCTTGCCGACGAACG
CGAACAGCGCTCCGCAGCGAGCCTCGTACCCCACACGCTCACGGATGAGACCCGAAAGCCGCT
CGAAGCCGTAGCGCATGTCCACCGGCTCCAGCGCGACGAACACCTGCACGCCCGCCGGAATCA
TCGCCCCGCTCCGCCGAGGGCACGGACCACCTCCGCCAGCAGCGCGGGTCGAACCCCGCGGC
GACGCGCACCCGCGCGCCGCCGACCTCGACGACGAGCTCCGCAGCGCTGCTCGTCACGGCGGG
CGCCTTCGGCACCAGGCGCAGAAAGCGCGGTGGCTCGGCCCGCGACAGCCGGCTCGACCAGCC
GTGCAGCGTCGAGGCCGCAAATCCGCGGCTCCGAGCGAACTCCTCCGCCGTTTCACCACTCTC
GCGCCACGCCCGAACGCGCTCGGACCACATCACTTCGGTCGCCTTCGTCCTTGTCATGCACGC
CATCATGAACTGGACAGCGCAGCCGGGGTGAGACGGCGCTTCGCGCAGCGCTTACGCAGAAGG
CGCGCCGCGCGCCATTGTCGGATGCGGTGCGCGACTTCGCCGCCGATCGGCTGTTGCTGGAAC
TGGGACAACCACTGGACGTAACGGCTGAAGCGAGCCAACGGCTCCAGCTCGCGCGGGGCGACC
TGTTCGGCGCCTACCAAGCGTTGGCCCAGCTCTGGATCTGCGGCGCCCTGGCCGAACCGCCGC
GACTGTATCCCGACGAACACCGCCGGCGCGTGCCGCTGCCGAGCTACCCCTTCGAGGGAAAGC
GGTTCTGGATCGAGGGCTCGCCGTTCGAAACCGCGCCCGCCGCCGGCGCCTCACCCCAACCCG
CCGATTCGGGGGACATTCTCAAGGGCGACCCGGCGGACTGGTACTATCGGCCGCGTTTCGAAG
CGGCGCCGCTCTTGCCCAGCCCGTTCGAGAGCGAACCCGGCGATTGGCTGGTGTTCGAAGATG
AGCTGGGGCTCGGCGCCTGGCTGAGCGAGACCTTGCGCGACAAGGGCGCGCGGGTCGCGACAG
TCGTTCGAGGCACCGAGTTCCGACGCCTGGCGTCACAGCGCTTCCAGCTTCGTCCCGATCGAC
GGGACGATTACCGGACCCTGCTGCACGAGTTGAAGGCGCAGGGCATCGCGCCGGTCCACCTGT
GCCACCTATGGAGCGTGACCGCCGCACCGGATGCCGAGCAGTTGCTCGACGTCAGCTTTCACA
GCCTGGTCCATTTGGCGGCCGCTTTGGGTTCGGTTGGCTACTTCCACGCCATGAAGTTGAACG
TGGTCGCCAACCGGCTATTCGACCCCGAGTCGCCCGAGCGCACCGAGCCCGCCAAGAGTCTGT

```
TGCTCGCGGTGACCAAAGTCCTGCCGCAAGAGGTGCCCAACGTTCGAACCCGCGCCATCAGCG
TGGACCTGGATCGCTCGTTCGACGCGGCGGCGCCCGCCTGGGCCGCCAGTTTGTTGGTTGAAT
GCGGCGCGCCCGTCGAGGAAACGGTGGTGACCTACCATGGCGCAGCCCGATGGCTGCGCCGCT
TCGATCGCGTTGCGGTGAATGGTCTCGGCCCGTTCCACCCCGATCAACCTGCGCCGCTGCTGC
GCGAGCGCGGCGTGTACCTGATCACCGGCGGCCTGGGCGGCGTGGCTGGCCAGTTGGCGCGCT
ACCTGGCGCGGGCCTGCCGGGCGCGGTTGGTGCTCACCGCGCGCCGGCCCCTGCCCGAGCGCG
ACCAGTGGGATCGGGAGTCGGCCGTGCTGTCATGGGACGACAAGACGCGCCAGCGCATCGAGC
TGGTGCGCGAGCTGGAGCGGCTGGGGGCCGAAGTATTGGTGGTGGCTGCCGATGTCGCCGACG
AAGCGGCCATGGCGCAGGCGATCGAGGCCTCACTGGCGCGATTCGACGCTTTGGACGGCTTGA
TCCACGGCGCCGGGATCGTGCGGGTCGCGTCGGGCCGCACGCCGATCGGGAGTATGACGCGGG
CCATGTGCGAGGAGCAGCTCCGCCCCAAGATGTTGGGCCTCGACGTCGTCGACCGCCTCCTGC
GCGATCGCCGGTTGGACTTCCGCATTGCCATCTCGTCGCTCGCCCCGATTCTCGGCGGCCTCG
GCCACGTCGCCTACGCCGCCGCCAACCTCTACATGGACGCGTTCGCGACGCGCGCCGCCGCCG
GCAACGCGCCTTGGATCGCGCTGAACCTGGCCGAGTGGGAATACGAGGGCCCGGCTACCTACG
ACGAGCGGGTGGGCCGTTCGCTCAAGCAGCTCGAGCTCACCAACGAGGAGGGTATCCGCGTCT
TCCAGACGGTGTTGGCCTTGGCCGCGCGCGGCCCGCTACAGCAGATCATTATTTCCACCGGCG
ACCTCCAGGCCCGCCTCGACAAATGGATTCACATCAAATCCCTGCATCGCCGACCGGGGCCGG
TCCAGCTCAGTCGCCGGACCGCGGCACCCCAGGGCGGTTTCGGCTCGGAGCGCGCCGCCTTCG
AGGCCGCCTTCGCTGACGCCTGGTGCGACTTCTTCGGGGTTGAAGAGGTCGACCCGAACAAAA
ACTTCTTCGATCTGGGCGCCAGCTCGCTCGACTTCATCCACCTCGTCAGTCGCTTCAGCAAGG
CCATCGAACAGCATGTACCGCTCGAGGCCCTGCTCGAACACTCCACCCTGCACGACCTCGCCG
CCCACCTCGCGGGCGACGCGAACACCGACGCCAGCGACGAAGCGCGCATTCGCCAACGGCTGC
AAGGCGCCAAGTCCGGCGACATCGCCATCATCGGCATGGCCGGCCGCTTCCCGCTCGCGCCCG
ACCTGGACACCTATTGGCGCAACCTGGTCGGAGGCATCGACGCGGTCAGCTTCTTCAGCGCCG
AGGAGTTGCGTGCTGCTGGCGTCACCGCGGCCGAGATCCACCACACCAACTACGTGCCGGCCA
AGGGGCGCTGCGCCGACCAGGACTTGTTCGATGCGGCCTTCTTCGAATACACTGCCAGCGACG
CCGAGCTGATGGACCCGCAAAATCGCGTGTTACACGAGGTCGTGTGGCACGCGCTGGAAGACG
CCTGTTTCGACTTCAACGGCGATCACGGCCAGGTCGGCCTGTTCGCGGGCGCCTCGCCGAACC
TGTGGTGGCAGTTCGTGGCCAGCTTTTCCGAGGCCGCCAAGACGCAGGGCATGTTCACCACCA
CCCTGCTCAACGACAAGGACTCGATCGCGACCCAGATTTCATACAAGCTCGGTCTAAAGGGCC
```

CCGCGGTCACCTTGTTCACCGGCTGTTCCACCTCGCTGGTAGCCGTTGACGCCGCCTGCCGCT
CGATCTGGTCCGGTCAATCGGACATGGCCGTGGCCGGCGCGGTCTCGCTGACTCTCCCCGATA
AGGCCGGCTACATCTACGAAAAGGGCATGCTCTTCTCGGCCGACGGCCATTGCCGGGCTTTCG
ACGCCAACGCCACCGGCATGGTCTTCGGCGACGGCGCCGGCGCGATCGTGCTCAAGCCGTTGG
ACGCGGCCCTGCGCGACGGCGACCCGATCCATGCGGTGATCAAGGGCTGCGCCACCAACAACG
ACGGCGACCGCAAAGCCGGCTACACGAGCGTCAGCGCCCAAGGCCAGGCCGAGGTGATCCGCT
CGGCCCAGATCCTGGCCGACGTGGCGCCCGAATCCATCAGCTACGTGGAAGCCCACGGTACCG
GCACCAAGTTGGGCGACTCGATCGAGATCAAGGCGTTGAAGCAAGCCTTCGCCAGCGACAAGA
ACGGATTTTGCGGCATCGGGTCGGTCAAGACCAACCTCGGTCACCTGATGGCGGCGGCGGGGA
TGGCCGGCCTGATCAAGACGGTTCTGGCGATGAAGCACCGCCAATTGCCGCCATCGCTGCACT
GCGACGAAGTGAACCCCGACCTGGAGTTGGAGCGCAGTCCGTTCTACATCAACACCCGCCTGC
GCGACTGGGTTGCACCGGGCGGGCCGCTGCGGCCGGCGTGAGTTCGTTCGGGATCGGCGGAA
CCAACGCTCACGTCATCCTGGAGGAGCCGCCGACGCGCGAGAGCGGCACGCGCATGCGCCACT
GGAAATTATTGATGCTGTCGGCGGCCAGCGAGGCGGCGCTCGACCGCCAGGCCGATAACCTGG
CCGACTACCTGGAGCGCCATCCCGAGGCCCACCTCAGCGACGTGGCCTATTCCCTCCAGACCG
GCCGGCGCGTTCTGGCCTGGCGGCGCACGGTCCTATGCGAGTACCGCGAGGACGCGGTGACCA
GTCTGCGCGAGCGACAGGCCAAGCGCGTCCAGACAAGTCGCGTCCGCTGGGACCACAAGGACG
TGGTCTTCATGTTTCCCGGTCAGGGCGCCCAGTACCTCAACATGGGCCGCGACTTATACGTCA
TGGAGCCGGTCTTCCGCGAGGTCATGGACCGCTGCTTCGAGTTGCTGGCCCCTTTGTGGTCCG
AGCATCCGCGCCAGATCCTTTATCCGGAGGGCGGGGTGTCGACCCTGCTCCACCGGACTGATT
ACACCCAGCCGATCGTGTTCTGCTTCGAGTACGCCCTCGCCCATTTGCTGCTCTCCTGGGGAT
TGAAGCCGGCCGCGACCATCGGCTACAGCTTCGGCGAGTACGTTTCTGCCTGCCTCGCCGGCG
TCTTCTCCCTGGAAGATGCGATCCGTCTGGTGACCGAGCGCGGTCGGCTGATGGCGGCTTTGC
CCGCGGGCGCCATGCTCAGCGTCCCGGTTCCCGAATGCGAGCTGCTGCGGCTGCTGGACGGCT
TCCACGCCCAATCGGCGGCCCATCTGGCGCTGGCCGTCGACAATGGCGCCTCCTGCATTGTGG
CCGGCGAGCAGGCCGCCATCTCGGCCTTCGAATCGATGCTTCGCAAGAAGCGTCTGTTGACCA
TGCGGGTCGCGGTCAGCCACGCCGCTCATTCGCAGGTCATGACCGGCGCGACCGACGCCCTGC
GCAGCATCCTGCGGAAGATCCCCCTCTCCGCGCCGACAATTCCCTTCATTTCCTGCGTCACCG
GCACCTGGATCACTGCACAGCAGGCTACGGATCGCGAGTATTGGGTGAACCACATGTGCGGGA
CGGTGCGGTTCGCGGCGGGTCTGACCGAGCTGGGTCAAAACCGCGAGGCGGTGTTCCTGGAAG

TAGGTCCGGGCCGCGACTTGACGTTGCTGGCCCACCGCATCCTGGCCGACAGCGCGGCCGTGT
TCGAGCTGGTCAAGGCGCCCGACGGCGGCGACGACGATGGGTTCCTCCTGCTGGATCGATTGG
CCAAGCTCTGGAGGCTGGGGATTTCGATTGACTGGGCCGGCTTCTACGCGGATGAGCGGCGGC
GGAAACTCTCGCTGCCGGGATATCCGTTCGAGCGGCGGCGCTTCTGGATCGAGGGCAACCCGC
TGGAGATCGCCGCCGGCAGGCCCAATGTCCAGGGGCCGCTGGTCAAGGCGTCGGACATCGGCG
CTTGGTTCTACGTGCCGCAATGGCGGCGGTCGGTGCTCGCCGAGCCGGGTACAACGGCGGCGG
GCGCCGCCGTCACGGCGGAGCAGGCACGCGTCGTGACCGAGCTACGGGCGGGATGCGCGTCGG
CCGGCTTGGGCAGCGGGGCCTGCGGACTGAATGGCGGTGCCCCGTCCGAGCGTCCGAAAGAAA
GTGTAGCGCCAGCCGGGTCGACCAGCGCAGCGGCGCAGACCGGCGCGGACTGCCCGACACCGA
CTGGGAGCCAGCGGCTGTGCCAAAGGACGGGGCCGAGCCGCGGCCGACCTGGCTTATTTTCG
CCGACGCCGGCGGATTGGCCGAATCTTTCGCCAAGCGGGTTCAGGCCCGCGGCGAGAAGCTTT
ACCTGGTGGCTTCCGGCTCGCGCTTCGAGCGCCTGGCCGAGACCCGCTTCCGCCTCGATCCCG
GGGCCAAGTCCGATCACCGCCTGCTTTTCAAGGCGCTCGACGAGGCCGACATCCTGCCGACCC
ACCTCCTCGACTTCCGCTCGCTTGACTGCGGCGGGCCCGACGCCGACCCCATGGACCAGGCCG
GCTTCTTCGGGCTGTTGCACCTGGTCCAGGCGATGGCAGAGGCCGGCTACAGCCATCCCATTC
GGCTGCTGATCGTCAGTTGCGGCGTCTACGATGTCACCGGTGCCGAACCGCTGCAGCCGGCGC
GGGCCACGATGATCGGACCGGCTCTGTGCATCCCGCAACAGTATCCGCACCTCGAAACGAGCC
ATGTGGATTTGGGCGTGGTCCATGCCGACGAGCTCCACGCCGCGCGCCAGCTCGACAGCCTAC
TTGCCGAATGCCTAAGTGCAACGGCCGAGCGCCAATTGGCGCTGCGCGGCCGACACCGCTGGC
TGCTGGACTACGAGCCAGTCCGCTTGCCGCCGCTCGACCCGGGCCGTCTGCCCTGGCGCCAGC
GCGGGGTCTACTTGATCACCGGCGGTTTGGGCGGGATCGGCCGCATCCTGGCCGAACACCTGG
CCCGCACGACCTCGGCTCGCCTGGTCCTAATCGGCCGCGAAACCCTGCCCGACCGCGACGACT
GGGACGCCTGGCTGAACCGCCCGCAACCGGTCGACGCCACCCACGAACGGCTGCTGCACAAGA
TCCGCGCGATTCGCGATCTGGAAGCGCTAGGCGCCGAAGTCCTGGTCCTCGCCGCCGACGTCG
CCAACGAAGCCGCCATGCGCGAGGCCTACGATCGCGCCGAATCCCACTTCGGCACAATCCACG
GGGTGATTCACGGCGCCGGCCTGATGGACGCGCAAAGCTTCTCACTGATCGACGCCCTCGACC
ACGACCTCTGCGCCCGCCAGTTCGAAGCAAAAATCCGCGGCGTCTGCGTGCTCGACCGCGTTC
TGGCCGACCGCACGCTCGACTTCTGTCTGCTGATGTCTTCCATCTCCACCGTGCTCGGCGGCC
TGGGCTATTTCGGTTACGCCGCGGCCAACGCCTTCCTCGACGCCTTCGCCCAGGCGCGCAGCC
GCGACGCCGCTTTCCCCTGGCTTAGCGTGGCCTGGAGCGATTGGAAGTACTGGACCGAGCGCA

AGATGGACAACGAGGTCGGCGCCGTCATCGACAGCCTCTCGATGGAACCCGCCGAGGGCTTCG
AAGCCGTCACCCGCGTCTTGGCTTGGGGCAAGGCGCCCCACATCGCCAACTCGCCCGGTGACC
TCGGTCGCCGCCGGGATCAATGGGTCAAACTGGCCAGCCTGAAATCGGCGCACTCCAGCGAGC
CCGAGCCGGCTAGGCATGGACGTCCGGCGCTCTCCAGCGAATGGGTCGCGCCGCGCAACGTGG
TCGAAGAGAAGCTGGTCGCCATTTTCGAGCAGGTGTTCGGCACTGCGGCACTGGGCATCGAGG
ACAACTTCTTTGAGTTGCGCGGCGACTCGCTCAAGGCGGTCATGACCGCGGCCCGTATTCAAA
AGGAGCTGAACGTGGAAGTGCCGCTGCCGACCTTCTTCCAGATGCCCACGGTCGCTGGCCTGG
CCCAGTTCGTGACGCAAGCCAAGCGCAGCGGCCGGGAGACGATTCGGCGCACCGCGCCGCGCC
CACATTACCCGCTCTCGGCTGCCCAGGGCCGCCATTACCTGCACTACCGCATGGACCCGCGTT
GTACCGCATACAACGATCCCTTCGCCAACCTGATCGAGGGTCCGCTGGACGTGGATCGCGTGG
AGCGCATCCTGCACACCCTCATCCTACGCCACGACTGCTTCCGCACCTCGTTCCACTTCCGCG
AGGGCGAGCCGGTCCAGGTGATTCACGATCGGGTGGACTTCAACCTGGCGCGGATTACCTGCG
CGCCCGAGGATTTGCCCGAACGGATGCGCGATTTCATCCGCTCCTTCGATCTGGAGCGACCGC
CCGCCATGCGCGCCGGCCTCTTCGTCACGGGGCCCGAGCGCCACGTGCTGCTAATCGATTTTC
ACCACATTATCACCGATGGCGTGTCGTTCGAGAACTTCGTCGGCGAGTTCGCGGCGCTCTACC
GCGGCGAGATCCTGCCCGAGCTGGAACTCGAGTACAAGGATTTCGCGGTGTGGCAGCATGAGA
ACCGGGGCCGCCGCGCCAACAGCGACCAGGCCCGCTACTGGACCGAGCAGTTGGCCAATGCGC
CCGGGCCGATCGAGCTAACCACCGATTTCCCCCGTCCCAGTCGACGCAGCTTCCGCGGCGACC
GCGTGCGGACCGTGCTTGATGCGGAGCTCGTTGCTCGACTCAAAGAGCACGCGGCGCGCCTCG
GCATCACCCTCTATAGCCTGCTGCTGGGCGGATTCTCGTTATTGCAGCACAAGCTCTCCGACT
CGCACGACATCGTCATCGGTTCGCCCGTCGCGGGCCGCACCCGGAGCGAACTCCAGGATCTGC
TGGGCGCGTTCGTCAACACCCTGCCGATGCGCCACCGCATCGACCCGACCCATACCGCACGGG
TCTTCTTGGAGCAGGTCCACCAGACAACCTTGGCGGCCCTCAGCTACCAGGAGCACCCTTTTG
ACGAAATGGTGGCGACGCTCGGGTTCGCCGCCGATCCGGCTCGCAACCCGATCTTCGACACGA
TGTTCTTGCTGCAGAACATGGCCATGGGTGCAACCACCATTCCCGGTCTGCGGCTCTCGCCTC
ACGACACTTTTCACCGCAAGGCATTGTGCGACCTGATGCTACAGGCGACCGAGTATGACTGCC
ACCTGGAGCTGGTGCTCGAGTTCGCCACCGACCTGTTCCGGCTGGAAACCGCGCAAGTCTTGC
TCGACCGCTACCGCCAAGTCTTGGAGTGGCTGTTGGCGTACCCCCATGAATCGATAGACGATT
TGACGCTCGCCGGCCACTTTCGCGAAGTCGAAGTGACGATGTCGGACGAGGGCGACTTTGATT
TCTCAGATTTCGAACCCCGCAACGTGAGAAACCTATGGCGCGCCTGAGCCGCACAGATCTCCA

ACTCGCCATTCACCAGCGCACCGTGGAGCGCGAATATTGGCGCGCTCTGTTCGAGCGCCATCC
GCAACGGTCCAGTTTGCCGGGGTGCTCACCGCCCCGATCGGCGACGAGTCGACCCGCGAGAC
CTTGTCATTCGTCCTCGACGAAGATCCCCTTCGGCTGAGTAATCGTTCGCCGCAACGCCTGCT
CACGGTGTTGGCGGCTGGCCTCGCGGCTTTCCTCCACCGCTGCGACGGCGCTGAGCGCTTCAC
CCTGGGGTTGGCCCTACCGCGCCAAGCCGATGACCATCACCCGATCCTCAACAGCTTGATCGC
GCTGGGGGTCGCGGTCGACTCGAGTACGACCTTCCGCGATCTGCTCTATGCGCTTCGATCCGA
ATACCACGAGGCGATGCGCCACGCCAACTTTCCGCTGGCGACCTGGTGGCGCGGCCTACCCGG
CGGAACGGCGCCGTTCGACGTCGCCCTCAGCCTGGACCCCTTCACAGACGGCGATTCGCTGGA
AGACCACGCGATCGGCGCGTTGTTCCGGTTCGCATTGGAGGGTGAGCGCCTCACCTGCCGATT
GCGATTCGACCCTGCGCGCTATGACCGTCCCGCGATCGAAAACCTCGCCGATCGTTTCGCCCG
CTTCCTCACGCGCCTGTGCCGGGACGCCTCCACCGTCATCCAGGCGCTGGACCTTTCGCTGCC
AAGCGATGAATCGGTGTGGCGCGTCACTGAAGGCGTGCGGCGCGGCTATTCGCAAGACCTGAC
GCTAGACCGCGCGTTCCGCCGCCAGGCCGCGCAAACGCCCGATCAGCCGGCGATCACGTTGAA
CGGGGACGTCCAGAGCTACGCCGAGGTCGACCGCCGCAGCGACGCGCTGGCCCGCCACCTCCG
TCGCCACGGCGTCGGTCCGGAAACGATTGTGGCCGTCAACGCCCGGCGCGGGCCTAATCAGCT
GACGGCCCTGCTCGCGGTCCATAAGGCCGGCGGCGCCTACCTGCCGATCGATGCCGAGGAGCC
GGCTGCCCGCCAGCAATTCAAGGTGCGCGACAGCGGGGCGCGGTTGGCACTGGAGCCGTCGCC
GGACCAGGCGCTGACCGTCACCGACCTGCCGCGGCTCTTCCTGGACGATGCCTCGCTCTTCGC
TGACGGCGGGCTCGATGTGCCGCGCGGCGCCGACTCGCTCAATCCGGCCTATGTGATGTACAC
GTCCGGCTCGACCGGACAGCCCAAGGGTGTGGTGGTTCCCCACCGCGGCGTGGTCAATCGTTT
GAATTGGGGGCAGTCCCGTTTCCCGCTGGACGAACGCGACCGAATCCTCCAAAAGACGCCGCT
GCTGTTCGACGTGTCGGTCTACGAGCTGTTCTGGGGCGCATGGAGCGGGGCCACCCTGGACAT
CCTCGAGCCCGGCGCCGAGCGCGACCCCGACGCAGTGGCCAGGGCCCTGGCCGAGCGCGCCAT
TACCGTATGCCATTTCGTGCCTTCGATGCTGCTCGTCTACTTGGAAGTCATGCGGCGGCACCA
TGCGCCGCCCGTGCCGGACCGGCTCCGTTACGTCTTCGTCAGTGGCGAGGCCCTCGAACCGGA
CCACCTCGCCGGGCTCCAGCAGATTGGTCGGCGCCTCGGCCGCACGATTCCCCTCGTTAATCT
GTATGGACCAACCGAGGCCTCGATCGAAGTCTCCTGCTTCGCCTGTCCCGCCGACCATGTGCC
GCGCCGGATCCCCATCGGGCAGCCGATCGACAACGTCGCACTGCACGTTCTCGACCGGCGCGG
CCGTCGCCAGCCGCCCTATCTTCCTGGCGAGCTGTTCCTGGCCGGCGACTGCCTGGCGCGCGG
CTACCTCAACCGTCCCGACCTGACCGCGCTCCACTTCGTGCCCAATCCCTTCGGCAACGGCGA

```
GCGCATGTACCACAGCGGCGACTTGGCGCTCGTGCGCGGCGACGGCCAAGTGGCGTTTCTCGG
CCGCCGTGACCACCAAATCAAAATCCGTGGTCAACGGGTCGAACTGGGCGAAATCGAGAGTCA
TTTGCGCGGGCTCGAAGGCATCGCCGCCGCCGTCGTCCAGGCCGAGTCGCAGCACCATGAAAC
CCTGCTGCACGCCTACGTCGTCACCAACGACGCGGGCCTCAATGCGGCCCGGCTGCGCGCCGC
CCTCGCTCAACATCTGCCCGAGTACATGATTCCCCAGCGCTTCTCGCGGCTGGCCGAGTTGCC
GCTGCTGGCGGCAGGCAAGATCGACCGCGCCGCCCTCGCGCAACGTGCAACGCCGCTCGCCAG
CGGCGCGCCCTTCGTGGAACCCAGCGGGCCCACCCAGCAGCGTATCGCAGAACTGTGGCGCCA
GGTCTTAGCGGTCGCCGAAGTCGGCGCCGAGGATCCCTTCTTCAGCATCGGCGGCAACTCGCT
CAATGTGCTCAAGCTCAGCGCCGCGCTGAGCGACGCCTTCGCGCGTGACATTCCCATGCCGGC
CCTGTTCCAATACGACACCATCGCCGCCCAGGCCTCCTGGCTCGACGGGCAGGTTGACGAACG
GGCCCAATCCGCCGCGCTCGACCGGCAGGCCGCCGAGGCGGCGCTGACCCTTCAAGAGACCGT
GGCCATTTTTGAGGGATTCGATGACGAACCATGACCATCACGAGGAGAGCAGCGGCCTGGAGA
TCGCCGTCATCAGCATGGCCTGCCGATTCCCGGGTGCTGCCGATTGCGACGCATTCTGGGAAA
ACCTGATCAACGGGACCTCCTCGATCACCCATTTCAGCGACGACGAGCTGATCGCGGCCGGCG
TTGACGCGCGCGACCTGACGCCGCAGTACGTGCGCGCGGCCGGCCAGATCGATGACGCCGAAC
GGTTCGACGCGGCCTTCTTTGGGTACTCCCAGCGTGAGGCCGAGCTGATGGACCCCCAGTTCC
GCCTGCTCCATGAATGCGCCTGGTCCTGTCTGGAACAGGCCGGCATCGATCCGCGCGTCGAAG
CCGCGCCGATCGGGCTGTATGCCGGCGCAGCCGACAACACCTACTGGAACGCGCTCTCGTCGC
TCGACCGGGGCTCGGCCGAATCGGAGCAATTCGCCGCCGAACAACTTTGCAACCGCGATTTTC
TGTGCACGCTGGTCGCCGCCGCGCTCAACCTGAAAGGCCCCGCGGTGGTGGTTCAAAGCGCCT
GTTCGACCTCGCTGTTGGCGGTCCACTCGGCCTGTCGTGCGCTCCTGACCGGCGAATGCCGAG
TGGCCTTGGCCGGTGGGGTGGCGCTGCGCTTCCCACGCCCGAGCGGTTATCGCTACGAACCTG
GCATGATCTTCTCGCCCGACGGGGTGTGCCGGCCGTTCGACGCGGGCGCTAACGGGACGGTGC
CCGGCGAAGGCGCGGGGCTGGTAGCGTTGAAGACGCTGAAACGTGCCCTCCAGGACGGCGACA
CGATCCACGCCGTGATTCGCGCGACCGCGGCAAACAACGATGGTGCCCGCAAGACCGGGTTCA
CCGCGCCCAGCGCCCACGGCCAAGCCGAAGTCATTCGCACGGCGCTGCGCCTGGCCCGGGTGC
CGGCCGAATCGATCGACTACGTCGAGGCCCACGGAACCGGCACGCCGCTAGGCGACCCGATCG
AGGTAGCCGGCTTGGTGGAGGCCTTCGCCAGCGAGAAGCGCGGCTATTGCCGGCTGGGCTCGG
TCAAATCCAACCTTGGTCATCTGGACACTGCTGCCGGCATCGCCGGCCTGATCAAGACCGTGC
TGGCGCTCGAGCACGCGCACATCCCCAAGTCCTGCCACGTCGCCACGCCCAACCCCGCGGCGC
```

GCCTACACAAGACGCCTTTCCGCATTGCCGCCGACGGGATGGCCTGGCCGCGGCGTATGGCGA
CGCCGCGGCGGGCGGCGGTGAGTTCGTTCGGCATCGGCGGCACCAACGTCCACGCGATTTTGG
AGGAGGCGCCGCCCCGCGCGCCCGAGCTGGCGGACGGGCGCAGTCAGGTGTTCGTCTTCTCCG
CCAAGGACGAGGCGGCGCTGGACCGTGCCCTTGCCAACTATGGTGCGGCCTTGGAGAAGCGCG
GCGACCTCGCGGCGGGCGCGGTGGCCTGGACGCTCCAAAACGGCCGGGCCGCATTCGAATGGC
GAGCCAGCGCGGTGGCATCCGACCTCGACGAATTGGCGGGCGCATTGCGCGGCGAGCGGCCCG
GCGCCGTCAAGAAAAACCGAATGGCGCGCGAGGATAAGCCGGTGGCGTTCTTATGTTCGGGGC
AGGGGAGCCAGTACCGTGGCATGGGCCACGACCTGTACCGCGAAGAGCCGCGTTTCCGGCACC
ACCTCGACGCCTGCCTCGCCATCCTCGCCGAACACAAGCCCGAGATCGACTGGCTGGCGTTGC
TGGGCTACCGCGACGAGGACGAGCCAACCGACCAGATCGGGACGTCCTCGCAGGGCCCGAGCC
GGTCAGCCGCATCGAACCCAGCGGAGCTCCTCGACAGCACCGAATTCGCCCAACCTTTGCTTT
TCTCCATGTCCTACGCGCTCGGTCGGCTGTGGCTCGACTGGGGCGTGCGACCCACGGCGATGA
TCGGGCACAGCCTGGGCGAGTACAGTGCTGCATGTATTGCAGATTTCTATGCACTCGATCAGG
TGCTGCCCTTCATTCTGACCCGCGGTCGAGTCATGGCGCAATTGCGGCGCGGCTCGATGTTGG
CCGTCAGCGGTGACAGCGTTCTGATGCGCGAGCTGATCGCCGATGCGCTCGATTTGGCGGCGA
TCAACGGCGCTGACCAATTTGTCTGGAGCGGGCCGAGCGAGGCTGTCCAAGCCGCGGGGGTCC
GACTGCGCGGCGCCGGCCTGCGTGCCACCGAGCTGAACACCTCACACGCGTTCCATTCAGCCA
TGATGGATCCCATTCTGGAGGAGCTAACGGTTGCCGGTTCGCGACTTCAGGTCGGTGTCGGGA
CGATTCCGGTCGTTTCATGCGTTACCGGAACCTGGTTGACGGCGAAGCAGCTGGCCGATCCGC
GCTACCACGCGCGTCACGCGCGCGAACCGGTGCGGTTCGCGGCGGGCCTAGCGACGCTGACAG
GGGAGGAGCCGCCGCTGATGCTCGAAGTGGGGCCGGGCTCGACCCTGGCGGCTTTGGCCCGCG
AGCATTCGAATGCCCGCCTCCCGGTCGTCACCAGCCTGCGCCACGCTCGCCAGGCGACGCCCG
ATCGCCAATACCTGCTCGAAACGCTCGGCTGCCTTTGGCGACACGGGGTTTCCGTCGATTGGG
GGGCCCATGCCGGACGTTCGCGACGCTTGGTTTCGCTGCCCGGCTATCCCTTTTCCGGCGCGG
TGCGCCGCTTAGCCGGCGACCCCCTCCGCCTGCTGGCCGGAGCCCGCGCCGTCGCCGCCCCGT
CGGGAACGCGCCAACTCAGCGCCGACGCGCGCGACCTCCCGAACACTCCGGAGCCGACATCCG
GCGCCGTGTCGGCGATCAAAGCGCCAATCGCCGCCGCCGATCCCGGCCTCTATCGCCTCTCCT
GGCGCCAGGCCGGAACGGCGCCGCTCGGTCCGCCCGATCTCGGTCCGCCCCGCGACTGGATCG
TCTTCGCCTCTGATTCTCACCTGCTCCAGGCGCTCAGGGCCAATCTCGGGACGCGCGCTCAGC
GGGTGACGCTGGTGACGCCGGGCCAGGAGTACGCAGCCGAGCCGTCCGGGTTTCGGCTGCGGC

CGGACCAGATCGACGATTACCGCGCCCTGTGGGCGGACTTGGCGCAAACCGGTATTGTGCCAC
GATACATCGCGTTCCTCGCCCCGTTCATGTACCGGGCGCGCATGGCGGGCGATGCCTCGACCC
TGGACGAAGTGCGCGAGGGCGGCTTCCTGCCCCTGACCCGCTTGATCCAGACTCGCCCGCCAG
GCGGACCGAGCGGACTTCTAAGCCTCACGATCGTCACCCCGGCCGCCCTGGCGCTGGGCGACG
AAGCGACGCGCCCGGAATGGGCAATCCTGCACGGGATGGTCGCCGGCTTAAGCCGCGATTATC
CCGAATGGCGCTTCGTCTCGATCGACGGCGGCGACCCATCCCCGCATCGGTGCGAAGGTCTGG
CCCGCTTGATCGCGCTTCATGCGGTCGACGAGGCTGGCCCGACCCGCTTGGCGCTGCGCGGCC
TTCACGCTTGGGTTCCACAGTGCGAGCACGTTCAGCCGGCCACCATCCCTGGGGCGGGTATGT
GGCGCGAGGGTGGTGTGTACATGATAACGGGCGGATTCGGCGGGATCGGTCTGGCGCTGGCCC
GCGCCCTGGCTCGAGAAGCTCGCGCCAAGCTGATCCTGGTCGGCCGAAACCTGCCCACCGCGC
CGATCGATCTCGAGGCTTGGGACGCGCCGCCGTTGATTCTCACCGCCGACGTCGCCGACGAAG
AGGCCATGCGCCGCGTCTTCGATGCCGCGCACGCCCGGTTCGGCGCCATCGACGGCATTCTTC
ACGCGGCCGGTGTCCCCGGTGGCAGCCTGTTCGCCAACCAATCGGACGCGGCCTTCGAAGACG
TGCTGCACGCCAAGGTTCGCGGTACCCTCGTGCTGCAAGGCCTGAGGGCAATCGATGCGCCGC
TGTTGCTGATGTCCTCGCTGGACGCCTGGCTTCCCGGTCCCGGTCAGACCGCCTATGCCGCCG
CCAACGCCTTCCTCGACGCCTTCGCCAGTCTGCGCCGGCGAGAGGGAGAGCCGGTGTACAGCG
TTGGCTGGGACAGTTGGTGCGAGGTGGGCATGGCTGCTCGGGTCGCTGCCCGATCGGCCGACG
AACGCGGCCGCCTGGCGCGCGAGGGGATCAGCCCTCGCCAGGGTTGGCAGGCTTTGAGCCGGG
CGCTCGCCCTCGACCCCCCCCACCTGATGATCTCGCGCACCGACCTGACCTCGCGCTGGCACA
GTCGATCCAGCCCTACGCCGGTCGCCTCGAGCGAACCCGAGGTGGCGCTGCCGCGCTGGACCG
CATCCGCCTGCCAAGCCGTCATCGAGCGTGTTTGGTGCGAGCACTTCGCCACCGCCGCCGTGC
CTCCCGATGGCAACTTTTTCGAGCTCGGCGCCAGTTCCTTCGACATCGTCCAGCTCAGCGCTC
GACTTCAACAACAGTTCGGCCGAGATGTCAGCCACACCGTGCTCTACAGTCATCCCACCGTCG
CCTTGCTGGCCGGCTACTTCGCCAATGACCCGACGCCGTCCGGTGCTGCTGCCGACGAACGCG
ACGAAGCGGTGCGTCGCGGCCGCGACCTCTTGAAGAGCCGCCGGCGAGGAGTATGACCGTGGA
GCACGAAACCGGATTCGAAATCGCCGTCATCGGGCTGGCTTGCCGCGTTCCCGGCGCTGCCGA
CGTGGCCGCCTTCTGGCGCAACCTGGTCGAGGCCAAGGAGAGCGTGCGCTTCTTCGAGGACCA
CGAGCTGCGGGCCGCCGGCGTGCCCGAGGAGATCTTGCGCCTGCCCAACTACGTGAAGGCCAA
GCCACTGCTCGCTGATGGCGAAGCTTTCGACGCGGACTTCTTCGGGTTCCATCCGCGCGAGGC
CGCCTACCTGGACCCGCAAGTTCGGCTCCTGCACGAATGTTGTTGGACCGCGCTGGAGGATGC

CGGCTACGATCCCGCGCAGTACGCCCTACCCGATCGGGTTGTTCGCGGGCGTCTCCAGCAATCT
CTCGTTCCTGTTCGACCGCATCGATCCGCGCGACTCCCCCCTGCAGAAGCGCTATGTGGCCGA
GCTGAACGCGGCCTCCTTCGCCACCCCAGATCGCCTACCGGCTCGATCTGAAGGGGCCGGCCAT
TTCGATTCAAACCGCCTGTTCGACGTCACTGGTGGCGATTCACCTGGCGGCGCAAAGCCTGAT
CGGCGGCGAGTGCCACATGGCCTTGGCCGGCGGAGCGACCTTGGAGGTCCCCAAAAAGCCCGG
CTATCTCTACCGCGAAGGCTACATCAACTCGCCGGACGGCCACTGCCGGGCCTTCGACGCCGA
CGCGGCCGGCACCATCTTCGGCGACGGCGTCGGCATCGTCCTGCTCAAACGCTACCGCGACGC
CCTACGCGACGGCGATCACGTGTACGCAGTGATCAAAGGCTCGGCGATCAACAGTGACGGCCA
TCGCAAGGTGTCCTACACGGCGCCGGGCAAGAGCGGTCAAGTGGCGGTGATCCGCGCTGCGCT
GGCGGCGGCCCAGGTAGAGCCGCAAACCATTCGCTTCGTCGAGGCCCACGGACCGGCACACT
CGCCGGCGATCCGATCGAGGTAGAGGCGTTGACGGAGGTCTTTGCCGAAGCGGGTCGCGGTAC
CTGCGCCCTGGGTTCGGTGAAGACCAACATCGGCCACTTGGATGTGGCGGCGGGCGTGGCCGG
TTTCATCAAGGCGGTCTTGGCGCTCGAGCGGCGCGTCCTCCCGCCCAGCCTTCACTTCGTCCG
GCCCAACCCGGCCATCGATTTCAACGGGCCCTTCTACGTTTGTCGCCAAATCGAGCGGTTGAC
GGAGAACGGGCGGTTGCGGGCCGGGGTGAGTTCCTTTGGCATTGGCGGCACCAATGCCCACGT
GATTCTGGAGGAAGCGCCGGCGCCGGAGGCGAGACTGCCGGCCGGGAGCCCGCCAGGCGCGAG
TCCGTTCCTGTTCCCGCTATCGGCCAAGACGCCGGATGCGCTGGCAGGCCGTTGCCACGACCT
TGCCGACCACCTGCGGGCGCACCCCGAGCTCCTCCTGGCCGATGTGGCCCTCACTCTGCAGAT
GGGGCGGGCGTCGTTCGCCTACCGCCATGTGGTCCAGGCTGCGACGGCGGAGGAGCTGATTCG
CGGTCTGGGAGCGTTCCGACAGGAGTCCATCCGCAAGAGGCGGAATCGAGTACAATGGGTGTT
GGCAGGCGAGGCGATGTCGCTTGACGCCGGTTTGCGGCTGTACGCCGATTGGCCGGTCTATCG
GGAGCGGGTCGACGTCTGTCTGGCGATCGTCGCCAAGCTGCGCCAAATCGACGGCCGGTCATT
CCTACATGAGTGGATCGAGCGACCGGCGCGAGGTTCCTGCCGAATGGTCGACGGCGCTGGCGTT
CATGTTCCACTGCGCGCTGGCGCAAGCCCTGAGCCAGGCCGGCCTGCACCCGCAGCGCATGTG
GAGCCGTGGGCTGGGCGGACAGGTCGGCGTGGTTTTGGCCGAATCCCTGTCGTTGGAACAAGC
GCTGGCGCTGGTGTTGTGCCAGACACCGGTTCCCGGCGATGCCACACCTCAGCGCGAACGCTT
GGTTCGGACACTGGAAGGCTGCCGGTTTCGTCCACCACGATTTTTGATTTCGGCAGACAGCTC
GGGTCGACCCCTGGACCTCGCCGAATTCGCTCATGTCGATTTTTGGTGCGGTGGCCAAAGCGC
CTCGCCCAATGAGGCGGAGCTGCGCTCATGGAGCGACGCCGCGCCCGAGCTGGTGACCTTGGC
GATCGGCCCATCCTTTCTCGAGGCCGCCTCCGGGACGGTGGGTCTGGCGATCGACCCCAAGCG

ACCGATGACCTGTGTTCAGCGCACGGTGGCCGCGTTGTGGGAATGGGGATGTGACGTGCGCTG
GGCTGCGTTCACCTCGTCGACCGGGCGTCGGGTTCCCCTGCCTACCTATCCCTTCGTGCGGGT
AATTCCCACGATCGGCGACCCCCTTCGCGGAGCAGGCGCGGAGGATGACTTGATTGCGGCGAG
CGCTTCCGCGTCGGCCGGATCGCCGCCCGAGCCGTCGGCAAACTCGGCAGCGGAACGCCCACG
CGCCCAGTCAAGCATCGCCTCGGCAACCACACCGGCTCCGTCTCATACGTCGGCCAGCGTGGC
CGTGGCCACCATTCTCGAAACCGTCCGTGCCTATTTCGGGTTCGCCGCCGTGCGTTCCACCGA
CGCCTTCTTCGAATTGGGCGCGTCCTCGCTGGATTTGGTCAACCTGGGCCAGCTCCTTTCCGA
TCGTCTCGGCCGCGAGGTTCCGACCCTGCTCCTCTACGACCACCCAACACCGGACCAGTTGGC
GCTGGCCCTGACATCCGCGGCGCTCAGCGCAGAGGCGCCGCCCTTAAGGGGCGGTCATCGCGC
ATCGACTTCCGGCACAGCCGCGAGCTCGGCCGCCTCCACCGCACCGACGTTCCCGGGGGACGC
TCACTCGCAGCCCAGCTTCGTTCGCGAGCAGGACATCGCCATCATCGGGATGGCCTTCCGGGG
ACCGGGCGCCGACGACCTGGACGCGTTCTGGAACAACCTGGTCGAAGGGGTCGAGTCGATCAC
CTTCTTCAGCGAGGACGAGCTGCTGGCGGCGGGCGTCCCCCGCGAACATCTGGCCTCGACGCG
CTACGTGCGGGCCAAGGGGGAACTGACTGGGATGATGGATTTCGAACCGGAATTTTTCGGTTA
TTCGGCGCGCGAGGCGGCGGTCATGGACCCGCAGTTCCGCGTGTTCCACGAATGCTCCTGGCA
CGCACTGGAGCACGGCGGCTACGATCCGACCCGATGCGCGGCATCGATTGGCGTCTACGCCGG
CGTGACCAACCACCTGCCTTGGCTGATGCGAACTTTGCCGCACCTGACCGAGGAGGAGCAATT
CGGCGCGCTGCTCCTCACCGACCGCGAGTTTTTCGCACCGCTGCTCTCCTACAAGGTCGGCCT
GCGCGGACCCGCTATTTCGCTGCAAACCGCCTGTTCGACGTCGTTGGTGGCGATCGGCACGGC
CTGTCGCGAATTGCGCGCGGGTGCCTGTCAGATGGCCCTAGCGGGCGGCGTGACGGCCAGCAT
CGAGCGCTGCGGCTACTTCCACCAAGAAGGCTACATCCTCTCGCCTGACGGCCACACGCGCAG
CTTCGACGCGGCGGCCGCCGGCACGGTCTTCGGCGACGGAGTCGGCATGGTGCTGCTGAAGCC
GCTGGCCCAAGCCTTGGCCGACGGCGACACGATCCACGCGGTGATCAAGGGAATCGGCATCAA
CAACGACGGCGCGCGCAAGGTCGGCTTCACCGCACCTAGCCGGGCCGGTCAGACCGAGGCGAT
TCGGGCCGCGCTGCGCGACGCCGGGGTGGCGTCGAACCGCGTCAGCTACGTGGAGGCGCATGG
AACCGCGACCAGAATGGGCGACCCGATCGAGGTCGAGGCCTTGACCCAAGCCTTTCGCGCCGA
AGCCGACGGTCCGCTTCCGCCCGGCTCCTGCCTACTCGGCTCGGTGAAGTCCAACGTGGGCCA
CCTGAACGCCGCGGCCGGCGTGGCTGGTCTGGTAAAAACCGTGCTGGCGCTCCAACACCGCCG
CCTGCCGACCAGCCTGTTCTACCAGTCGCCCAATCCACACATCGACTTTGCGGCGAGTCCGTT
CCGCGTGAACGGCCAGACTTCGGATTGGGTCGCGCCAGAGGGGACGCGGTTGCTGGCGGGAGT

GAGTTCGTTCGGTATCGGGGGAACCAACGCCCACCTGATCGTCGAGGAGGCGCCGAAAGCGCT
ACCGACGACAGCGGCACCTCTGTCGACGGAGCCGAATGACCTCGACGCGGGCGACGCCGACGG
GCTAGTGCTGCCGATCTCGGCCCGCACGCCGACCGCCCTGGCGCACATCGCGACCAACCTCGC
CAATCACCTGGAACGACATCCGACCATCGCCCTGGCCGACGTCGCCCTGACCCTTCAGCTGGG
CCGTCGCCAATGGCCCCATCGCCACAGCCTGATCTGCCGGAATCGAACGGAGGCGATCAAGCT
GCTGCGCGCCGTCGTCCACTCCGCGGAGGTGCCGCCAGCTCAGGCGCCGGTCTCGGATGCGCC
GCGCTGTGTTTTTCTTTTTCCCGGCCAGGGCGCCCAATACCCGAGCATGGCCCGCGACCTGGT
TCGAAACTGTCCCGACTTCGCCCTGCACCTGGACCCCTGCCTCGACCAGTTGGCCGAACTGCT
TCCCGAAGATCCGCGTTGCATCCTGTTCGGCGATGGCCCCGCCGATCGGCTCGACCAGACGGC
CTACACTCAGCCGCTGCTCTTCTCCGTGTCCTACGCCTTGGCGCGCTGGTTGGGCGATTTCGG
CATTCGCCCCGATGCGATGATCGGCCACAGCCTGGGCGAATACGTGGCGGCCTGCTTGGCCGG
GCTTTTCTCGCTGAGCGATGCCCTGCTGCTGGTGAGTGAACGCGGCCGCCTGATGGGCTCGGC
CGCGCGCGGAGCGATGCTGGCCGTCCCCTTGCCCGAATGGGAACTGGAGGAACGCCTGGAGCT
TCTGGCCGACGACCGAATCAGCATCGCGGCGGTCAACACCGCCGAGAGCTGCGTCATCGCGGG
ACCCAGCGAGGCGATCGAGCGCTGCGCCCAGCGCTGGGCCGCGCAAGGCCTGACCTGTACGCC
GCTGCGCACGTCCCACGCCTTCCACTCCGCGATGATGGAGCCGATTGTCGAACCCTTCGGCCA
TGTCTTGGCACGGGTCACCTTCGCGCCGCCGCGCGCGCGCTGGATCTCGAACCTCGACGGCAA
GCCGATCGATTCCGCGGCGGTGATGCAGCCCGACTATTGGGTGCGCCACCTGCGCCAACCGGT
CCGCTTTCACGAGGGACTCAGTCACCTGTTGGCCGAGGACACCCATGCTTGGGTCGAAGTGGG
TCCCGGCCGAACCCTGTCCTCCTTCGTCCGCCGCCACCCGGCCTACCGTCACCAGCCAATCGT
CAACCCCATGCGCCATGCAGTCGAGTCGACGGGCGACGTGCGCCGGTGGCGCCAAGCGCTGGG
CGAACTATGGCGGGCCGGCATGCCGGTCGCCTGGGAGCGGCAGCGGCGCGGCCGGCATGCCGG
ACGACGTGTGCCGCTGCCGGGCTACCCCTTCGAGCGGCGGCCCTTCGCGGCCCGAAGACCGGT
GGAGCTGGCGCAGCCCGCGCCCAAGGCGGAGCTGGTGAAAAACCCCGATCCCGCGCGGTGGCT
GTACCGCCGCGTCTGGCGCCCTGCCCAGGCTGCGGCCGGCGGACTGGCGGTGCAGGCGACCGT
TCTGGTCTTCGGCGACGGGTCCGAGCTGTGCCGCGCGGCGGTCGCTCAGGTGCAGCGCCAGGG
GCTGAAGTGCGTCTCGATCACCGCGGGCCGCCAATTCGCGCGGGAGAGCGACATGCGCTTCAC
GCTTGACCCCGCTGATCCGCGCCAGCTCGACCAGCTCTTCGCGGCCCTCGATGGCTCAGGCTC
GCGGCCGCGGTACGTCCTGCACCTGCTGACCCTGAACCCGCCCCCGGATGCCTCGGCGATCAT
CGCTCACAGCTACTACAGCCCGATGGCCTTGGCTCATGCCTTGGGCGCCCACGAGATCGCGCC

TGTCTCGATCACCGTCGTCACCGCCGGGGTCGTCGCCGTCGCGGACGAAGCGATTCGCGAGCC
GCTGCAGGCGCTGATCGTGGGCCCGTGCCTGGTCATCCCGCAGGAGTTTCCCGGGCTCAGCGT
TCGGCTGCTGGACGTCAACGTCGACGATCCGGCACCGCGTCTGGCGGAGCGGCTCGTGGCCGA
GCTCTCGGGCACGGATCACATGGTGGCGCTGCGCGGCGGCGAGCGCCTAGTGGCCGATGTCGA
TCAAGTCGATGGCCTCGGTGTGGGGATCGCCAAGGTGCCCTTGCGCCGCGAGGGCCACTACCT
GATTCTCGGCGGCCTGGGCGATATCGGCTACCACTGTGCCCGCTATCTGGCCCAAACCTACCG
CGCCAAGCTGACGCTGACCGCGCGTTCGTCACTCCCGCCGCGCGCGTCGTGGGAGCGAATGCT
GCGCGAGGGAAACCTGGATTCCCGGCAGCGCACGCGCATCGAGCGCGTGTTGTCGCTAGAGGC
GTGCGGGGCCGAAGTCCAGACGGCTGCGGTCGACTTGGGCGATCGCCATCGCTTGGCCGATGT
GTTCCGCGAAGCACGGGGCCGATTCGGCGCCATCGCGGGCGTGATTCACTCGGCGGGGATTCC
GGGACACGTCCACTCGATCGACGAGCTGGTGCGCGTCCGCGACGAAGCCCAATTCACCGCGAA
GGTTCGAGGGCTGCACCACCTGGCCGAGGTCGTCGATCCGCTGAACCTCGACTTTTGTCTGCT
GTTCTCCTCGCTCTCGACCGTCCTCGGCGGGCTCGGCTACGGCGCCTATGCAGCGGCCAACGC
CTACATGGACAGCTTCGCCCGCCGCCACGATCGGCCGGACGAATGTCGTTGGATCGCGGTCAA
CTGGGACGCCTGGCTGTTCGAAGCCAAGACGTCGTCGGTCGGCGCCGAATTGGCGCGCCTGGC
GATCGTGCCCGAGGACGCTCCGGCCCTGTTCGCGCGGGTGCTAGAGCGACTTCCGCAATCGTT
CATCGTGTCCACCGCCGACCTTCGGGCCCGCATCGACACTTGGATCCGGGACAAGAACCGCGT
CCCGCCCGCCGAGATCCGAGCGGTTCAACCGCGACCGGACCTGAGCCAGGCGTACGCCCCGCC
GATCGGCCCGCTGGAGATTCAACTCTGCGGGCTGGTCTCCGCCTATTGCCGGTTCGACCGGAT
CGGGCGGGACGATTCCTTCTTCGAAATCGGCCTCAGCTCGTTCGACTTGATCCAGCTCAGCTC
GCGCATTCACCGCATCACCGGCAAGGATCTCAATACGACCCAACTGTTCAGCTACCCCACCGT
GCGCGCCTTGGCGCTCTTCCTCGGCGGCGAACCGGAGGGGCTCGCGGCGGAGGAGCCCGCCAT
GGAGAACCTGTGGCTGCAACGAAGCGATGCGACCCTCGATGAGTGAGACCGAGGTCGCCGACT
GCGGCGCTACCGACCGCGGTCGAGGATTTTCCGCGCAGCGATCCGGGACGACTCGCTGAAGAA
GCGCGATAGAAGAACGGAATCGTGTATGAAATACGAAACCACCGGATTGGAATTGGCCGTCAT
CGGTCTCGCTTGCCGCTTTCCAGGCTCACCCGATCCCGAACAGTTCTGGTCGAATCTGCGCGC
AGGTCGCTCCGGAATCCGCCATTTCAGCGATGCCGAGCTGAGCCACATCCCCGCATCCCTGCG
TCACCATCCGCATTACGTCAAGGCCAAAGGCGCGCTGGACCACGCCGATTTCGAACCAGCCTT
CTTCGGCTACTCGCCCAAAGAGGCCGAGGTGATGGACCCTCAATTCCGGCTGCTCCATGAGTG
CTGCTGGGAGGCGCTGGAGTCAGGCGGCTATGCGCCGAGCCAATTCGCGGGTCGGATCGGCTT

GTTCGCGGCGGCGGCCTTCAACGACGGATGGATCGCCGGTACCCTCGACCGGCTGCGCACCGG
CGTGGGTTTGAGCTCCCTGGAAACCGCGTTCTTGACCCTGCGCGATTACCTGACCACCCAGAT
CTCCTATCGGCTCGATCTGCGGGGCCCCAGCCTGCTTGTCCAAACCGCCTGCTCGTCGTCGCT
GGTGGCGGTCCAGCTCGCCCAGCAGGCGCTGATCTCCGGCGAATGCGCCCTGGCCTTGGCTGG
CGGCGTGTGCGCGACCGATCCGCTGCATTCGGGATACCTCTATGAACCCGGCAACATCTACGC
GCGCGACGGCGTCTGCCGACCGTTCGACGAGGCAGGCGCCGGTACGGTCTTCGGCGACGGGTG
CGGCATGGTCCTGCTCAAGCGGCTGAGCGACGCCCAGCGCGACGGCGATACGATCTGGGCGGT
CATTCGCGGGGCGGGCGTGAACAACGACGGGCACCACAAGGTTGGCTACACGGCTCCTGGCAC
GAGGGGCCAGGTGGCTTTGCTTAAAAGTGTTTATCGCGCGAGCCGGGTCGACCCGGCGACGCT
CGGCTACCTGGAGGCCCATGGCACCGGCACCGCGCTCGGCGATCCAATCGAGGTCGAGGCGCT
TACCCAGGCCTTCGCCAGCAAACGTCGCGGCACCTGCGGCTTGGGCTCGGTCAAGGGCAACCT
GGGTCACCTCAACACGGCGGCCGGCATCGCTGGACTGATCAAGGTGGTGCTGGCGCTGAAACA
TCGCGAAGTGCCACCCACCCTCAATCTGCGCCGTCCCAATCCGAAAATCCGCTTCGACGAGAC
GCCGTTTTTCCCAGTCGTCGAGTTGCAACCCTGGCCAAGCGGGACCGGCCCCTTGCGAGCCGG
CGTGAGCTCCTTCGGCATCGGCGGTACGAACGCCCACGTCATCCTCGAGGAGGCACCGCCGAC
GGCCAACCCGGCGCCACACGGCAGATTCCGACTGTTGCCGCTTTCGGCCAAGACACCGGCTGC
GCTCGAAGCGAAGCGCCGCGATCTGGCCGGCTTCCTCGAACGCCACCCGGAGACCTCCTTGGC
CGACCTCGCCTTTACCCTGCAACGCGGCCGCGAGGTCTTCAGTCACCGCGCCTGCCTCGCCGT
GGAGACCTTAACGTCCGCGCGCACGCGGCTGAGCGGCGAGTCGTCGAGCACTTGCGTGGTGGG
CCCCGCGCCCAGCGCCATATTTCTGTTCCCTGGTCAAGGCAGCCAGCTCGCCGGGATGGGCCG
CGGTCTGTATCACCATTTCGAGCCGTTCCGCACGGCCGTCGATGCCTGTCTGCGCGAGCTGGA
GCCAGGACTGCGGCAAGCGCTCAGCGCCCATTTCGATCCGAATCGCGGCGCGGACCCACCCGA
TTCGACGACCTTCGTCCAACCCTTGTTGTTCCTCGTCGAGTACGGGGTGACCGAGTGGCTACG
CTGCTTGGGTGTGCGGCCAACAATGGTGTTGGGTCACAGCTCTGGCGAGTATGCCGCAGCCTG
CGTCGCGGGCGTTCTGTCGCCGTCCGCGGCGGTCTCGCTGCTGGCCGAGCGCGAGCGGCTGCT
GCGCGACCTGCCAGCCGGCGCCATGCTCGGCGTCCCGCTGGCCGCCGAGGCGCTCGAGGCGAT
GTTGCCCGACGCTCTCGATCTGGCGGCGATCAACGGCTGTCAGCTTTGCGCCGTGTCCGGGCC
GGTCGCGGCGGTCCACGCCTTCAAGGCCCAACTGGAAGCCGCCGGACATCACGCCCGCCTGTT
GCACACCGATCGCGCCTTCCACTCGCGGCTGGTAGCACCGGTGCTTGACCGGTTCCAGGCAGC
CGTTCAACACGTGGAGCTGCGGCGGCCGCAAGTACCTTACCTCTCGACCGTCAGCGGGCGATT

GGAGGCGGATGGGCCGGCGAACCCGCACTACTGGGTGCGTCACCTGCGCGACACGGTGCGGTT
TGGTCCAGCCCTGGAGGCGCTGCCGCCGGTGGATTCCTTCGTGTGCATCGAGGTGGGACCAGG
CTCGGCCTTGAGCACCATGGCGCGCGAAACGTTGGGTTCCCAGGCGCGACTGATTTCGTTGCT
GCCGCGGCCGCGAACGGGGCAAATCGAGCCCGGTCCGGTATTCGAACGACTGGCGGCGCTTTG
GCGCAGCGGGTTGACATTGGATTGGTCTAAATTGACGGGCGGCGAAGAGGGTCATCGAATTCC
CTTGCCAGTCTACCCGTTTCAGCGCAGCCATCTGTCGAGCTCCCTGGCGGCGGGCCACACGCC
TTCGTCGCGGCCTGCAGTCGAATCAGGCGCCATCCTTGCCGAGCGATCCGCAGGGGAAAACGC
TGAAACCCGGGATTGCCCGCTGCCAACCGCCACGCTCGAGCCCAAGGCGGTCGCTCCGGCCCC
ACTCGAGGCTACCGACGCCGCAGGTACTCGCGAGCGACTGGCCGAACTTTGGCGCGAGTTGCT
AGGGTTGACCTCGATTGGGCCCGACGACCATTTCTTCGACCTGGGCGGCCACTCGCTGACCGC
CACGCGGCTGCGCGCCCTGATTCACCAGCGGTTCGATGTCGATCTCGGGCTCGACGAAATCTT
CGCTCATTCGCGTCTCTCCCAGCTGGCCGCCCGTATCGAGGCGGCGGCCAAGAGCCGATTTTC
CTCCATTCCCAGCGCGCCGGACCAGGACGACTATCCCTTGTCATCCGCCCAGCAGCGGATTCA
CAGCATCGTCACGAGGGCCGAGGTCGGCACTGCTTATAATTTTCCGATCGTCCTCGAGCTGCA
GGGCGCTCTGGATCGAGTGCGATTCGAGGCGACGTTCGCGGCATTGTTCCGGCGTCATGAGGG
GTTCCGCACCCGCTTTGTGATGCGCGATGGCGGGCCGCGCCAGCGCATTGTACCGGACGTGGC
GTTTCGCCTGCCGCTCACCCAGGTCGAGCCAGAGCAGGTTCCCGGGCGCATCGAGGCCTTCAT
CCGTCCCTTCGATTTGGAACGCGCGCCGCTGTTCCGCGCGGAGCTGTTGCAGTTGGCCGAGCA
GCGCCATCTGCTACTTTTCGACATGCACAACTTAATTGCCGACGGTATCTCGCTCAACCTGTT
CGTCGCCGATTTCGCGGCCCTGTACCATGGTCGTCCGCTGGCGCCGCTGAAACTCCGCTATCG
CGACTATGCCGTTTGGCAAGAGGCGCGGCTGGCCTCCGATGACCTGCGCAGCCAGCGCGAATG
GTGGCACCGGCGGCTTTCGCCGCCGGTCGCCACGCTGGCGCTCCCTCCCGATTTCCCGCGTCC
GGCGGTGCGCCGCTACAAGGGCCGTAATGTGGTGTTCCACCTGGACCGGGAGATCCGCGACCG
CCTGGTGGCCCTGGCTCGAACCCAGGGGGTCACCATGAACGTGATGATGCTGGCGCTCTGGGC
TGCGCTGCTGCATCGCGAAACCGGCCAATCGGAGCTGGTGGTCGGATCGCTGCTCGGCGGGCG
GCCGCACAGCGAGCTGCATCCCGTGATCGGGCTCTTCACCAACTTTTTGCCCTTGCGGTTGGC
GGTCGAGGGATCGACCCGCTTCGATCGCTTCCTTGCCGCTTGCCACCAGGTGTTTCTCGAAGC
CTATCAGCGCCAGGACTATCCGTTCCACTTGTTAGTCCAGGAACTCGTGCCGGTCAGGGACCC
GTCGCGGTCGCCGCTGTTCCAGACCTCGCTCGTCTACCACAACGAAATTGACGGCAAGACCAA
GCTGGAATTGGAAGGGCTGAAAGTCGAAGTGGTTCCCTTCGAAAAGGGTGTGGCGAGGCTGGA

TTTGAAGCTGGATGTGACACCTTTTTCCGACCGACTCGAATGTGTTTTGCAATACGACTTGGA
TCTGTTCTGCGAGGAGACGATGCGCGGCCTGATCGCGCGGTTCCAGGCGTTGGTGGCGGGGCT
TGTCGCCGATCCGGCGCAATCGCTCGCCGCCGCGAGCGTTTCCGGGAAGCGGGCGCTGCGCGC
GGGCGTGGCCACGGCAAGCGAATCGTCGCCGCAGTCACTGCCGCCGCAACCATCGACGGCGTA
CGCCACTCCCTCACCGCAGTCACCGTCGCCGGTAGTCCTGACGGGACCCGCCGACCTGCCCGC
GATCTTGGCGGCCTACGTGGGGCAGAACCCCCATCCGTTCGCGATCCATCGGGGTCTCATTTT
GGAGGCGCCGCTGGGGTTGCGAGCGCTGCGGTCGGCGCTGGACGCAGTGCTCGGAGAACACAC
CCATTGGCGCAGCGTGCGTGCGGGCGATCGCGCGCGGCGCGTGGATAAGTTGGAATTGACCAG
CCTGGTGCGGCTCGACGACCTGCGCGGGTTGGTCAATCCTCAGGCGAATGCCTTCACCCTGGC
TTGGCGCGATCTGGCGATGCCGTTCGGGGAGGGGCGTCCCCTGTGGCGACTCCGCCTGGCGTG
GTCGGCTCCATCGCGCTGGTTGCTATTGCTGACGGTTCATCCATTGATCGGCGACAACGGCAC
GGTCGACCTCTTTCTGGCGGCACTCGCCGATCACCTGCGCCGCGCGTCCGCTTTTCCCGTAGC
ACCGCTCGATGAGGCCGAGCTGGAGGCGGAGCTGAAGTGGGGAGAGGAAGGGGAGGGCCTCGG
GCTGACCGCGATCGCGCCGGTCCTGGGCCAATTGCGCGAAAGTCGGCTGAGTCCTGTGGCCCA
GATGTGGCTGGACGAGGTCTGTCGCCGCCACGACCTCACCCCGCTAGAGGTCTTGGCGGCCCG
GCTCCTCGATTGGACACGAAGCCACGGTCACGGGTCGATCGCTTTGTGGACGCCGCTGCCCGA
GGACCATCCGCTTCGCGATGAAGGCCGCTGCCTCCAGGTTCGCCTGCTGGAGGGGCCGCCGTC
GCAGCGAGGAGCGGGCGATCCAAGCTGGCTCGAGCAAATCGCCTTGAGACGGGGTACCCCTGC
AACGGAGGTCGTTTGCCCTACTCCGACCCAACGGGCAGCCATCGACCTCGCGCTGGCCTGGCT
GCCGCAGCCGCCTCTTCACGGTTTGGTCGGAACCGTTCAGCCGTGGCCGGAATCTCCATTGGT
CTGTCCGTTTCCCCTCAATCTCGCGTTCCGGCCAAGCCATCCAATTGCCTACGCGCTCAAGCA
CGAGGCCACGCTCGCGGTCACGGCACGGGCGCGCGATCTGATGCGTTTCCTCGACGGCTTGGG
CCCGGAAAGCTGAAGATTAGCATAAGCGCCCGGCCAAGGGCATCCTAGGATGACGCAAGCCTC
GGCCGCGTCGACGTCCCAGGTCGCGCCGGAGGTCACCCCCGGCCGAAAGGACGACGATGACGA
TCAAATCCGAGATGTCGGCCGTTGCTCACTCTGCGGAGAGCGGCTTCCGCGCTGGGCCACGCG
TGGGCGGCGCGATGAAGCGGGGCCGGACGCCGGAGCAGGCCGGCGTGAAGCTGCTCCGCGCCC
CGGTGAAGCGGAAGTGGCTGCCCCCGGCGCCCGTCCTGCGCCTGAGCGAGCGGCGTATCCCGG
AGGTGTGGGCAGGCTACCGCGCGAGCGCGGGATGACCCGAGCCCCGCCCGCCGGCGCGACCAT
GACGCCGCCCCACGGGGCGAGTCGTCCGGCGCGCCGGCGCGCGTCGGGGCTTCCGCCGCCGGG
CGGGCAGGTGCAGGATGGTCGGGCATGGTGACGCGTCCGACGTCCGACGGCATCGAGGACGAG

CTCGCGCCGTTCCCCCCGGTCCTGCGCGGCTGGCTCATCGAGGGCGAGCTCGGCCGCGGCGGG
ATGGGGCGGGTGTTCCGGGCGCGGCACCCGAAGACGCGGGCGCGGGCGGCGATCAAGGTGCTG
CTCGGCGACTACGCCCGCCGGCCGGACGTGGTGGCCCGCTTCCGGCAGGAGGCGATCGCCGTC
AACATCATCAACCACCCGGGAATCGTCCGCGTCTTCGACTCCGGCGAGCTCGAGGACGGCTCG
CCCTACATCGTGATGGAGTACCTGGACGGCCGGGGGCTGCGCGACTGGGTGCAGGCCGTGCCG
CCCGCGGAGCGGCCGCGGCAGGTCGTGCGGCTCGGCTACCAGATCGCCTCGGCCATGGCCGCG
GCGCACGCGTCCAAGGTCGTCCACCGCGATCTGAAGCCGGAGAACATCATGGTGGTCGAGGAC
GAGCTCGCGCCCGGGGGCAGCCGCGTCAAGATCCTCGATTTCGGCATCGCGAAGGTCCTCTGG
GGAGGTCTGCCCGAGGTGCTGGAGCTCGAGGGGAGAGGCTCCCTCGCGCCCGCGTCCGCGTCC
ACGATCCGCACCGAGCTCTCGACGCGGCCGGCGCCGACGGTGGGCGCCACGACCGGCCCAGAG
AGCCCGCTGGGCGCGAGCGCCACGCCAGAGAGCGCCCTGGGCGCGAGCGCCACGCCAGAGAGC
GCCCTGGGCGCGAGCGCCACGCCAGAGAGCGAGGCCCACGAGGAAGACGCGCTCCGGAGCCTC
CCCGTCGTGACCAGCGGCAGGCCCGCGATCCACCCCGCGCCGGTCGAGATCCCGCCCGAGGCG
GTCTCCTCCGCGGCGTCGCGCGGGTCGCGCGCGTCGATCGAGCCAGGCGCGCCCGCGCCGCAG
AGCGAGGGCGCGGGACAGCCCACGATGCCGTTCACGCAAGAGGGCGTGTGGGGCCTCGGGACG
AGGAGCTACATGGCGCCGGAGCAGGAGCGCCACTCCGGGAGCGTGGACGTGAAGGCGGATGTC
TACTCGCTCGGCGTCATCCTCTATGAGCTGCTCGAGGGGCGGACGCCCGACGCGCCGAGCGCC
GCGTGGCCGCCCCCGATGAGCGCCGCCACGCCGCCCGATCTCGTCGCCCTCGTCCACCGGGTT
CTGGCGTTCGATCCCGATGCGCGGCCGCGCATGGCGGAGGTGGCGAGCGCGCTTCACCGGCTC
GGCCGGGCGAAGAAGGAGCTCGACGAGGCGCTCTCGAGGTGGGTCGTCGGCGGAGGGGCGCCG
GGGCTCTTGCCGTGCGGCTATGCTCTTCTCGAACTGGTCCTCCTGGGCCCTGGGAACTTATAC
GATTCTTTCCAGCCTGTAAGTGCATTTTTCTTTCAATATCGTCCTCTCTTCATATACGAGGTG
AGTTCTCTGAGGTCCTCCTATAAGTCTGGGGTGTCCTATTCGGCCTCTTACTTGTTACTTCGC
CTTCTTAGGAGTTTTTCCTTAATTTTGCCCTCTTACATTCCCGTATTCATTCTAACTGGGCCC
TATCTCATTCGCTAATACGTTTCTGTATTGTGTACATCTCCTATCATGTGTCAATACTTGTTT
CTGTTTATCATTATTCTTATTGTTTACGCTCTTATTTCATTCATAGTATAACATTAGTTTACT
GATTATCGCACTTGAATTCGCG

or its complementary strand,

(b) DNA-sequences which hybridise under stringent conditions to regions of DNA-sequences according to (a) encoding proteins or to fragments of said DNA-sequences,

(c) DNA-sequences which hybridise to the DNA-sequences according to (a) and (b) because of a degeneration of the genetic code,

(d) allele variations and mutants resulting by substitution, insertion or deletion of nucleotides or inversion of nucleotide segments of DNA-sequences according to (a) to (c), wherein the variations and mutants offer isofunctional expression products.

8. DNA sequence according to claim 6 selected from the following

(a) open reading frames:

	Nucleotide Position	
ORF1	1666 - 1	Seq ID No 3
ORF2	1605 - 3338	Seq ID No 4
ORF3	6100 - 3398	Seq ID No 5
ORF4	7110 - 6374	Seq ID No 6
ORF5	9590 - 8433	Seq ID No 7
ORF6	11393 - 9855	Seq ID No 8
ORF7	13656 - 12712	Seq ID No 9
ORF8	15374 - 18984	Seq ID No 10
ORF9	20003 - 27889	Seq ID No 11
ORF10	28251 - 29402	Seq ID No 12
ORF11	31720 - 30401	Seq ID No 13

ORF12	31982	-	32932	Seq ID No 14
ORF13	33128	-	33613	Seq ID No 15
ORF14	33661	-	34007	Seq ID No 16
ORF15	35611	-	35255	Seq ID No 17
ORF16	37856	-	35730	Seq ID No 18

or DNA sequences complementary to said open reading frames,

(b) DNA-sequences which hybridise under stringent conditions to regions of DNA sequences according to (a) encoding proteins or to fragments of said DNA sequences,

(c) DNA-sequences which hybridise to the DNA-sequences according to (a) and (b) because of a degeneration of the genetic code,

(d) allele variations and mutants resulting by substitution, insertion or deletion of nucleotides or inversion of nucleotide segments of DNA-sequences according to (a) to (c), wherein the variations and mutants offer isofunctional expression products,

and peptide sequences corresponding to said open reading frames

SEQ ID No 19 (>ORF1)

VDPEREAVTLGLAFNRAQGRTYARGPEARAEYIGTAMRAADVIEDRFEIERLAVSGGMGDVYR
ARDRVSGQAVALKVLQGASANDLRRFAREAEALVTLRLPGVVQYVAHGVTGAGRPYLAMEWLD
GVTLEERLAGAPLTLAESVALAARVATTLGAIHWLGVVHRDLKPSNLMLVGGAVERVTLLDFG
IARHLRLAPTLTSPGAVLGTPGYIAPEQVRGDAPVDARDVFALGCVLFQCLAGRPPFLGNSAL
ALLMRVVLEEPPRLGELRDGIPEPLERLVARMLAKNAGERPRDGAAAAAELAAVAGEGLSIGA
SAVAAPAAPGEAITTAERKVMCVILAEDGGAEAGATLSEDDGAARAEALRDIAARHGGRLDRL

QARWWLVALSGAESPTDLATRAAHCALALRAALGGVPVSVATGLAEVEARLPVGELVDRVAQL
IAGRDGLSPPEIRLDDATASLLASRFETVQGPGGCWLRGPKEEPDAVPRLLGKPTPCVGRERE
LSQLATEWRHCVDEPSANAVVVVGAPGLGKSRLAWEFLRTLEQREGAAI

SEQ ID No 20 (>ORF2)

VRPCARLNASPSVTASRSGSTAAGSVHASTSACVEQPATGRTQPASPRWPPGAAALRLTSAMP
RWFNTAGPCNPADHYMLPAEERLPAVRDLVDRKAYFVLHAPRQIGKTTSLRTLAQDLTAEGRY
VAVLVSAEVGAPFSDDPGAAELAMLAEWRGTAGAQLPADLRPPPFPDAPAGQRIGAALRAWAQ
AAPRPLVVFLDEADALRDATLVSLLRQIRSGYPDRPRDFPHALALVGLRDVRDYKVASVDSGR
LGTSSPFNIKVESLTLRNFTRDEVATLYAQHTAETGQVFRPDAVDRAFELTQGQPWLANALAR
QLVEVLVKDRAQPITSANVDRAKEILIERQDTHLDSLVDRLREPRIRAVIEPMLAGTALPSVP
PDDLRFAIDLGLVRMTAEGGLDVANPIYREIIVRELAFPIRASLPQIKATWLTQDGRLDADRL
LDAFLSFWRQHGEPLLGAAPYHEIAPHLVVMAFLHRVVNGGGTVEREYAIGRGRMDLCVRYAG
ETLAIELKVWRDGRPDPVAEGLAQLDEYLAGLGLDRGWLILFDQRSGQPPIAERTRRERALSP
AGREVAVIRA

SEQ ID No 21 (>ORF3)

VTIKKTFRSIDPATLPKHFDSPVAELRLADLWEADGTYRYDPSRPREETFVVDTPPPTASGSL
HIGHVFSYTHTDVVVRQRRMRGFNIFYPMGWDDNGLPTERRVQNYFHVRTDVRTPYERGLTLP
QAAPETIKKEPPRIVSRPNFIELCHKVTREDEQVFKALFRRVGLSVDWRNEYATIDDHCRRTA
QLSFLDLHEKGHLYSVFAPTMWDVDFQTAVAQAEVEDRPQSGAFHDIAFAVEGTAEELVIATT
RPELLAACVGVTAHPEDPRYQHLFGKTALTPIFRAPVPIFPSPLVDREKGTGILMVCTFGDAT
DVIWWREQKLPLRQMLGKNGRVLPVTFGEGAWESRDPAAANAAYAPLQGRGVKQARAAVVELL
RREEHAAAPGRGPALRGEPRPIERAVKFYERGDQPLEFVPTRQWFVRLADKKAELLEYGDKIK
WHPDFMRLRYRNWTEGLQGDWCISRQRYFGVQFPVWYPLDAEGNPDHSRPLLATREMLPVDPT
VDVPPGYEASQRDQPGGFTAESDVFDTWFTSSLTPQISSHWGDDPARHARLFPADLRPQAHDI
IRTWAFYTIAKAMLHESSVPWHHVAISGWILDPDRKKMSKSKGNVVTPMHLLDTYSSDAVRYW
SASARLGTDTAFDEKVLKIGKRLVTKIWNASKYVLSQSAEVHPISEELDRALLHKLSAVVDDA
TRSFDEHEFAAALERTEDFFWRWFTDAYLELAKARARGEGGAGEAARGSAVAALRLGLSVLLR
LFAPVLPYITDEVWRWVYAEETGDTSIHRAKWPSAADFAAVAAPSDPGLLDLAAAAMAAVNKR

KSELGASVGRVVTDLALGANAATLARLKPALGDVLTAVRAGAHALVRPELADGEVLVVRCELE
PAAAAAAGAGGAAASEE

SEQ ID No 22 (>ORF4)

MIHAEPFEARLVAARPLSPFVRELSFERADGRSFLFEAGQWVNLVLPLPGGEVKRAYSIASAP
DGSPRFDLAVTLVQGGAGSEHLHRLEPGATLRAIGPHGLFTRDPGDSAPSLFVATGTGITPLR
SMLRASLRAGLAAPHLWILFGARFEEDVIYRDELEALARGSDRIRYEITLSRGGPSWAGRRGY
VQAHVPELYRELAEKSGDPAPHVFICGLDRMVSSVRELARGELGVHRKHVHVERYD

SEQ ID No 23 (>ORF5)

MKSLPSDRAARLAQSDIRTMTLACAKVHGINMSQGVCDTPVPSVILQAVKEAMDRGCNTYSRF
DGIVELRHAIAAKLARHNGIAADPETDITVSAGATGAFQATCMALLNPGDEVLLFEPFYAYHA
QAILAVEAVPRYVTARSLSWNVDGDELERAITPKTKAIVVNSPGNPSGKVFGRMELEQIADLA
CHHDLMVITDEIYEYFIFDGREHVSVASLPRMSERTITIGGYSKTFSITGWRIGYSVADARWA
KAIGAMSDLLYVCAPTPLQHGVAAGIRGLPRSFYTGLAQGYERKRDRFCRALEKAGLPPCVPQ
GTYYVLADVSRLPGRTGRERAIYLLDETGVAGVPGDAFFEGTQGSRFMRFCFAKTDEDLEEAC
QRIEQLA

SEQ ID No 24 (>ORF6)

VSDPRKERLGDMDLEEFRRIGMRIIDWAADYLGHPDRYPVFPAIRPGDVKGRLAPTPPVEPEP
MDAVLTDFEQIILPGITHWNHPRFFAYFANTASGPGILGELLAACLNVNVMLWRTSPAATELE
ELVLSWLRQMLDLDAGLHGAIMDTASTASMVAIAAARDSAEPTIRLRGMAGQRRMRLYASEQA
HSSIEKAAITLGIGQEGVRKIPTDPAFRMVPEALRAAVVEDLGAGLRPFCVAATVGTTSTTSV
DPIPAIVSVCREHGLWLHVDAAYAGMAAIVPEHRDVLAGCEGADSLVVNPHKWLFTPMDCSVL
YVRDADRLKRAFSLVPEYLRTEGDVTNYMDWGIQLGRRFRALKLWMIVRYFGHEGLAARIREH
LRLGQQLAQWVDADPDWERLAPTPFSTVCFRMRPSALACIMRSADEAERESIERELDRLNEAL
LDEVNKSGRVFLSHTRLHGRYTIRVAIGNIRSDEVAVREAWECLRAAGARLCADERFVSCSRS
ADEGRGKS

SEQ ID No 25 (>ORF7)

MRREEPVLEAFYERYCAAPRETSYHVELPVDVELHQEAAPALPQARSLELAGRVALVTGSSRG
IGKAIALRLAEQGADVAVNYHSNKDAAEQTAAEIRALGRRTMVVQADVTRPNAAAELFSSVEA
QLGPIDILVNNVGDFFFKPLAAMTDDEWRNVMDSNLSSVHYLCRAAVARMRQRKSGRIINIGL
SPTYAIRGAPNVAAYSIAKTGVLILTRSLATEEAPHGILVNCVSPGLIDNGYLPPAQKEWMER
RVPMGRLGRASEVADAVAFLASDRASYVSGANIAVAGGWDWTDRGTEHDRRVDLFIGHEEP

SEQ ID No 26 (>ORF8)

MSGRFPGARNVEELWQKLRAGVECVVTFTEAEALAAGVSREMLANPSYVRRGAPLDGVELFDA
SFFGFSPREAESMDPQQRIFLEVAWEALERAGYDPDAHSGPIGVFAGSAPSGYHSLAQSDPEI
LGALGHYQLTLNNDKDYLTTHASYKLNLRGPSVCVQTSCSTSLVAVVMACQSLLNHECDMALA
GGVGIHAHQRRGYLYQENGISSPDGHCRAFDVAAKGTVGGSGIGIVVLKRLADALADGDHVHA
VIRGAAINNDGSSKIGYTAPSVQGQAEVIGMAQALAGVEPDDISYIEAHGTGTPLGDPIEIAA
LTRVFRAKTARRQFCAIGSLKTNLGHLDAAAGVASLIKTVMALEHRELPPSLHFERPNPKLEL
ESSPFYVNTRLTPWHAARGPRRAGVSSFGIGGTNAHVVLEEAPAPPPSGPSRRWQLLTLAARS
EAGLARATADMIEHLDRHSGTSIADVTYTSHVGRRAWPFRRAVVGESAADLRAALASEGSPRS
ISSCQAARERPVVFLFPGQGAQHLFMARELYEVEPIFRQSLDRCAELLRGPLGLDLRQVLYPA
EGQRDDAEQELGRTAIAQPALFAIELSLAKLWMAWGIVPQAMIGHSVGEFAAACLAGIFREED
ALRLVAERGRLMQQMPPGAMLAVPLAEPELAPYLSDDISLAAINGPALSVVAGPIEAIDALAA
ELLDHGLSCRRLHTRHAFHSKMMAPVVDAFTRCVSAVERRPPSGHFLSTLTGGWISPEAATIP
AYWARQLVEPVRFAQAVRQLLSESTWLWLELGPGQTLSPLVRQQARADGGQVVVASLPRAKDA
GADHLAVIEALGRVWSAGGTVDWKRFHEGEARRRVLLPTYPFERQRYWASPRHTSAPPEAIIK
PLLAKNPNVADWFFLPAWRRSDPPVSFDAQAVTTRRSTWLVFIGDEGLGAALVEGLARRGHEV
VAVVTGERFEQTGTQRYTIDPAANGDVASLFARLEIEGRMPDRIVHAFCTSPADGARIERGAA
LEIERRLGFDSLLLLAQVIAAQRHPKPLMLGVITTRAHSVIGTEIIEPLRALVLGPCRVIPQE
IPHVSCRNIDIDLPGEGGRAEIAARLIADLERESPDSVVAYRGGRRWVESIELTDVGRRSAGA
APRLRQRGAYLITGGLGGIGLVAAELLAREAHARLILVGRTGLPARQGWDDWLAAHGAGDATS
RKILRIRALEEAGAEVKIAAADVSDFNAMRSVIEEARTRFGRIDGVIHSAGIASGGMIQLRTP
MAAWRVMAPKVGGTLVLDALLRDERPDFLLICSSLASLVGGATQIDYCAANAFLDAYAQSREG
EEGCRVISVQWDTWSDVGMAVDFKLPADLQEGRRESLKRGISSSEGAEVLGRILSAGMSGPLA

ICTSDLPAYKQSVTTRRSQHEQTPAARPMHSRPTTTGAYVAPETETERRIAAIWQDLLGLEQV
GANDDFLQLGGHSLLATQVLSRVLQTLKVGISLPQFFDAPTVAGLSRLVDAARAEGAGPVAPA
IGRVERDAYRIKPPAAEQAARTKP

SEQ ID No 27 (>ORF9)

MEPVGGVDMNQPAKQQETCVFPTSFAQRRLWFLDQLEPGSAVYNMPASFRTRGPYDVDSLVRS
VNEIVRRHESLRTTVDVIDGEPVQVIAPSLRIEVPVVDLSEIDEPEREAEARRLMAEESRRPF
DLTRGPLLRAKLLRLGEADHVLILTMHHIVSDGWSMDVLFKELSTLYAAFHEGRPSPLPELPI
QYADFAVWQRELLQGEVLESHLGYWREHLRGAPTLLELPMDRPRPPAQTFRGSQRAFRLPLSL
QQAVQALSRQEGATPFMTLLTAFSVLLSRYARQSDLVVGTPIANRTRAELEGLIGFFVNMLAL
RIDLGGDPSFRELLGRVREVTLGAYAHQDLPFERLVEELSPGRSPSHSPLFQVSFTLQNTPMD
ATNRADIASGGAPLVEMKAAKFDLILELSESPQGLLGTFEYNTDLFDAGTIERMAGHLEVLLS
SAVAAPDRPIAELPLMGAEERSRVLVEWNSTAALYPEDHCMHELFEQQVERSPEATAVLLQQQ
TLTYRELNMRANQLAHHLRSLGVGPEVRVGLYLERSIETVVAILGVLKAGGAYVPLDPTYPSE
RLGLMMADAAPSVLLTQASLLSKLPPHGDATLVQLDALHEALSRLPHHTPRSGVTAQNLAYVM
YTSGSTGRPKGVLVEHRGLCNLPTVQAKLYGIAPGDRLLQFAPLCFDTSFCEIALALLSGATL
VMGTADELLPGPPLVELLKKHAVTAMLLAPTVLAALPEQQSAALPLRVLTMAGEACPAELVKR
WKAPGRRLFNSYGPTETTIWASSAADLSDERIPPIGRPIANTQIYVLDEALEPVPIGVPGEIF
IGGVGVARGYHGRPDLTAERFVPDPFGQTKGARLYRTGDRARWLPDGNLEFLGRNDEQVKVRG
VRIELEEIRAALLKHPAVAQAVAVVREDTPGDKRLVAYVVGRGGARVTAAELRQSVSERLPAT
MVPSSFVALDALPLTPNGKVDRRALPEPEQSAGGEDHVAPRNAVEEELARIWASVLRLERVGV
HDNFFEIGGDSILSIQIVVRAQQAGLRLTPRQMFQHQTIAELSTVARAVEAVHVEQDPVTGPA
PLTPVQRWWLEQEAAEPHHFNQSIFLEVRERLDESALEQAIAHLIDHHDALRLRLARDERGAH
QVFAAPGGSTPFQRVDLGALPSAEQISAMEKAASEAQASLDLAAGPVVRAVLFDLGEVAPQRL
LVIAHHIAVDSVSWRILLDDLFGAYEQARRGEAVRLPPKTTSVKRWAELLTEHAGSEAVKAEL
GYWLDSSRRTVAPLPVDRRAGEDVWGSARHIVVSLTPEQTEQLLREVPQAYRTRIDDALLTAF
AQAIARWTGSPAVLLDLEGHGREELAGVDLTRTVGWFTAMYPILLRVDAADPGEALKSIKEQL
RAVPGRGLGYGLLRYLRSDTIAEVRALPQAELCFNYLGQLDQAIPEAAPFRPAREYQGSERSP
GAHRAHLIEVNASIANGRLYATWTYSERRHEPETIERVAASFVTALRALIAHCTLPEVGGNTP
SDFDKVRLRQETIDALDAIDAGPGPSARGSRIEDVYPLSPLQEGILFHTLYATDYTAYVEQFH

WTLEGDFDAEAFTRALQDVVARHAALRTSFAWERLDAPLQIVRTGAVLPVEHQDLRGLAAEEQ
TAHISRYVEAERQRRFDLRKAPLMRAGLLRLRKDAWCLVETIHHLILDGWSTQILLKEVFTLY
EAHRGHRGHLALELEQPRPYGDYIGWLAKQDQVRTAAFWRRELEGFSAPTPLGVDRAVPHDDG
GPRFGWRRIALSGDDAARLAAFARQHQLTMSTLVQGAWALLLSRYSGDPDVLFGMTVSGRSAP
IPGIERMTGLFINTIPVRVREPADASVLAWLKALQEHEAELLEHEHSPLVEVQAHSDVPRGTP
LFESLVVFENYPVQVIFEAPPVEGPTRAEEGLRMIDAQYISDPPYPLTVVAAFHGTLYLNIGY
ERRRFDDQAVERMIGHVTTLLRGFVQRPETSVRDLPLLTAEEERTQLHAWNATAAPYPEGHCM
HELFEQQVERSPEATAVLLQQQTLTYRELNIRANQLAHHLRSLGVGPEVRVGLCLERSIETVV
AILGVLKAGGVYVPLDPTYPSERLGLMMEDAAPSVLLTQTSLLSKLPPHGDATLVQLDALHEA
LSRLPHHTPRSGVTAQNLAYVMYTSGSTGRPKGVLVEHRGLCNLPTVQAKLYAIAPSDRLLQF
APLCFDTSFCEIALALLSGATLVMGTADELLPGPPLVELLKKHAVTAMLLAPSVLAALPEQQS
AALPLRVLAMAGEACPAELVKRWKAPGRRLFNSYGPTETTIWASSAADLSDERIPPIGRPIAN
TQIYVLDEALEPVPIGVPGEIFIGGVGVARGYHGRPDLTAERFVPDPFGQTKGARLYRTGDRA
RWLPDGNLEFLGRNDEQVKVRGIRIELEEIRAALLKHPAVAQAVAVVREDAPGDKRLVAYVVG
RGGARLTAAELRQSVSERLPATMVPSSFVALDALPLTPNGKVDRRALPEPERSAGGEDHVAPR
NAIEEELTRIWADVLGAKRVGVHDNFFDLGGHSLLLVRVHDRLGQRFDRPPSMVDLFTYPTVA
SLARFLGERANGKQSPREAAADVTERGRRRLEARARRAKAIRGPT

SEQ ID No 28 (>ORF10)
MKHNIGWLLPAALATLAFVPACSPNHGEDAPSVTSAESGAAPSADCVALGAKLQAALDGAAAA
QKAPGAAAAVQSGDCVWRGATGVSDLVASTPTKPGDLFRIGSITKTFVSTLILMLRAEGRLSL
DDAVSKYVKGIPAGDQMTLRQILGHTSGLFDYTYSPALGQMIEVDPTRAFAPAELIALATAEA
PYFAPGAGFRYSNTNYIVAGLVAEAVSGGTLAGLLRTRILDPVGLAHTYLDGAEPPVQGLIRG
YGDYGAGLVDITDQLSPTEAWAAGALVSNVDDLNRFFALLISHELLSSDELQDMTTWTPTMWP
HEPGYGLGLIERDSALGSLNGHCGIIWGFQSASYGVPGRGDAITALINRSDGDAARLVDELAK
VVKER.

SEQ ID No 29 (>ORF11)
MSIDRAVLEQLDRVGGRLAEGKALKLLEDIAWPREVEERFFAAGEDRLPEVEYRVDRDGLARR
VAELRELLGAIDGDAPALGWLRDNVRAQIQAAELLEAAGTRAFSARSQELYGGARSRFFGGSL

RNIDLAEHLTERLRVHGWDEASDPEEEPLDAGALRDMLAARVAGRAPRLDLEITVDPRVTAKV
VAGMSRVRIRPEATFAAWEAEGLWHHEVETHALTAHNGAAQPRCAFLRSGGPRTTRTQEGLAI
FAELYSRSLSIGRLTRLAERVRLVDMAEQGASFLDLYRHLRERGAERRDAYFDAQRVCRGGLV
EGGAPFTKDACYLAGLLEVYAFLAAVLRGGLRDEVELLVCGRIALDDIAVLAELRAAGVLERP
RYLPGWLRAWQTLLPYFAFTSFMDGIDLGPVERHFQELLRVAADARPAGEGRRRRGRPREG

SEQ ID No 30 (>ORF12)
MSESVAQLEEHRAALTGHCYRMLGSVVDADDAVQETMVRAWRSLDKFDGRSSLRTWLYRIATN
VCIDLRADRARRARPIEEGPVGTVDDALETRPRTHWLEPVPDAHALPADIDAAERAMLRQSIR
LAFVAALQHLPPKQRAALLLTEVLGWSAAEVADSLNTSVAAINSALQRARATLASRDLGDARP
SLPEPQSALLDRYVNAFERYDVDALTALLHQDATLSMPPFTLWLRGHESIRAWLVGPGAGCRG
SRLIPTAASGSPAFAQYRPAPEGGHRAWALIVLDVAGDRIVSMTSFLDTETLFPRFGLPLDLP
A

SEQ ID No 31 (>ORF13)
VTIASIDHRDQDLMTGPQAKAPARAAAPDAAPSRRAVWAGRVLSGLATLFLTFDAAVKVLKLF
PAEASTAELGFPAHLVPTLGYLQIACLVAYLIPRTAVLGAILWTGYLGGAIAIHVRVENPLFS
HTLFPIYVAAFLWAGLWLRDRRVRALTASPSSQGR

SEQ ID No 32 (>ORF14)
MTTKNPRKLFVNLSVRDLKRSMEFFSKLGFEFNPQFTDEKAACMVVSEEAYVMLLVESFFKTF
MKKEICSTSTHTEGLFALSCSSRAEVDDMVKKAVAAGGSHAMDPQDHGFMYGWSFYDVDGHHW
EVMWMDPKAIQP

SEQ ID No 33 (>ORF15)
MTPSERLDATFAALADPTRRAILARLASGEASVTELAKPFAMSQPAISKHLKVLERAGLISRG
RDAQRRPCRIEAKPLEDASGWLDNYRRFWEGSYERLDDLLEELKERESKGERSKR

SEQ ID No 34 (>ORF16)
VAPASAPAAGGRDAAPFLDEAAQWLRGEQAPASRPAGEGPAGRLPGRVLVADDNADMREYALR
LLVAEGWTVEAVADGRAALERARAHPPDLVLTDVMMPRLDGFGLLRALRADDRTRGVAVVMLS
ARAGEEARVDSLEAGADDFLVKPFSAKELLARVRIHVELARRRREAEGQRQYLNDLFMQAPGP
IAILRGPEHVFEVVNPLYQRLVGGRSLVGEPIRAALPELEGQGIWELLDAVVRTGEPIVGKEL
PVRLDRRGDGTTEEVFFNFVYQPMRDRDGAVEGVFVFAFDVTDQVRARRRVEALVEALKLADQ
RKDEFLAMLAHELRNPMASISLSLTLLDDADGDGPASARYREIARRQMGHLVRLVDDLLDVSR
ITRGTVELRLEDVDLAAVVQSAAAAVRPAVEARRHDVSLSVGPGDFGMRADATRLEQVVTNLL
TNAAKYTPPGGSISVRLTREAAVGAPEAVLRVRDTGRGIPAAMLEKVFDLFTQVDQTIDRSTG
GLGLGLTLVRRLLELHGGSVAAASAGPGQGSEFTVRLPLGPGAAPQPAPSAGPPPPREGPPPA
QRDEPPPPPAQRAEAPEAAADRRRVLVVEDAEDVRRVMRAYIEALGHEVTVAVDGLEGVKKLL
ELRPEVAFVDIGLPGIDGYEVARRARAAPGGEALYLVALSGYGGPDDQARSRRAGFDLHLTKP
VVGATLQDVLTAPRT

9. DNA sequence according to claim 7 selected from the following

(a) open reading frames, and peptide sequences corresponding to said open reading frames:

pEPOcos6_ORF1 sequences:

(1) nucleotide sequence

Seq ID No 35 (>pEPOcos6_ORF1.seq)
GGATCACCTGCGGCGCGATCGCCGACCTCGTGCTGGTGTTCGGCTCGCTGGATGAGAAGCCGG
CGGCGCTACTGATAGAGACGGCGACGCCCGGGCTGCGGGTGGAGCGGTTGCGGGAGATGCTCG
GCTTTCGGGCGGCCCACCTGGCGAAGCTGTCCTTCGACGGTTGCGAGGTCCCCGAGGCTCAGC
TGATTGGCCGGCCCGGCTTTGCGCTGATGTATCTGGCCCCCTACGCCCTGGATTTCGGTCGGG
TCAGCGTCGCCTGGGCCTGCCTGGGCATGATCCGCGCTTGCCTGGAGACCTGCGCACAGCACA
TCCTCACCCGCCGCACCTTCGGCCACCTGCTAGCCGATCACGGCATGATCCAAACCCTGATCA

CCAACCTGGGGATTCACCACCAGGCGACGCTGCTCCACACGCTGCAGGCCTGCCGCGCCAGGG
ATCGCGGCGACGTGACCGCCTCCGAGGCCACCCTCGCCGCCAAATACCTCGCGTCGCGGACGG
CGGTCCAGGAGACGACCAACGCGGTCCAGATCATGGGCGCGCTGGGCTGCGACGAGGAGGGCG
CGATCGCCCGCCACTTCCGCGACGCCAAGACGACCGAAATCATCGAAGGCAGCAACCAGATCA
TCGAGGCGCTGCTGGCCAAGAACATCGCCCGCGCCGGTCGCGACAACTATCGCCGCTTCCTCG
ATGCGGAAGTCGAGCCCGGTCGGGCCGGAGGCGCACCA

(2) peptide sequence

Seq ID No 36 (>pEPOcos6_ORF1.pep)
ITCGAIADLVLVFGSLDEKPAALLIETATPGLRVERLREMLGFRAAHLAKLSFDGCEVPEAQL
IGRPGFALMYLAPYALDFGRVSVAWACLGMIRACLETCAQHILTRRTFGHLLADHGMIQTLIT
NLGIHHQATLLHTLQACRARDRGDVTASEATLAAKYLASRTAVQETTNAVQIMGALGCDEEGA
IARHFRDAKTTEIIEGSNQIIEALLAKNIARAGRDNYRRFLDAEVEPGRAGGAP*

pEPOcos6_ORF2 sequences:

(1) nucleotide sequence

Seq ID No 37 (>pEPOcos6_ORF2.seq)
ATGACGAGCGCGGTCCCGACGCGTCAAACCAGCCTGCTCGACGACTTCGAGCGCGTCGCCGAC
GTCGATCCAGAGCGGATCGCCGTCCACGCGAGCGAGACGAGCCTGCGCTATGGCGACATGAAT
GCGCGCGCCAACCGCATTGCCCACGGGCTACGGGCGCGCGGATCGGGCCCAATCAAATCGTG
GCGGTGGCGATGGCCCGCACGCCCGAGCTGATGATCGTGCTGTACGGCATCCTCAAGGCCGGC
GCGGCCTACATGCCCATCGCCCGCGACGCGCCGCCGCTGCGCCGCGATCATATGCTGCGCGAG
AGCCAGGCTGCTCTGATGATCGCCGACGAAGAGATCGCGGGACTCGCGGCCCGGGTGCTGACG
CCGGCCGACCCGTTCTTCGCGGCCATGCCGGACCACAACCCCGAGCCGCGTCACGACCCGACC
GACCTGATTTACGTCATCTACACCTCGGGCTCGACCGGCCAGCCCAAGGGCGTGGCCATGGAG
CACCGCGCCGTGTGGAATCGCCTGACTTGGATGCAGGCCCAGTATCCAATCGACACGCAGGAC
GTGATCCTCCAAAAGACGCCGATCGTCTTCGACGTGTCGGTCTGGGAGCTGTTCTGGTGGCCG
CTGGCCGGCGCCTCGGTGGCCCTGCTGCCGCAATCCATGGAGAAGTTCCCCTGGGCGATATCG
GCGACGGTGGCGCGGTGCGGGGTGACGGTGATGCATTTCGTACCATCGATGCTGATGGCCTTC

CTTCAGGTGGTGGCGGGCCGGCCCGAGATGGCGGACCAGATGAAGGGCCTGCGCTACGTCTTC
TGCAGCGGCGAGGCCCTGGCGCCGGCCCACGTGTCAGCCTTTCAGGAGCACATCAACCGAGCG
GGCAGCATCAGCTTGACCAACCTCTATGGACCCACCGAGGCGGCGGTCGACGTCAGCTACTTC
GACTGCCCGCCCGGCGCGTCACTCGCGCGGGTGCCGATCGGACGAGCGATCACCGGCATCCAG
CTGCTGGTCATGCGCGACGGCGTGCCTCAGCCGCCCGGCGTCGAGGGTGAGCTCGCCATCGGC
GGCGTTGGTTTGGCGCGCGGCTACATCTCACGGCCAGACCTGACCGCCGACCGGTTCGTGCCG
CATCCAGGCGGCGACGGCCAGCGGCTCTACCGCACCGGCGATCTGGTGCGCAGGGACGCGGAC
GGCGAGCTGGTCTTCCTGGGGCGCATCGACCATCAGGTGAAAATTCGCGGTCTGCGCATCGAG
CCCGGGGAAATCGAGGCCCAGATCAGCGCCCATCCCGATGTGGCCGACTGCGCGCTGATTATC
GAGCAGGACTCGGAAACCCTGCCCAAGCTGACCGCCTACATTGTCGTGGCGCGACCGGGCTTG
ACCCGGAAGGCGCTGCTACAGTTCCTGGGCGCGCGGCTGCCCGACTACATGCTCCCGAACCGC
TTCCTGACCCTCACGGAGCTGCCCGTGACCGCCAACGGTAAGCGCGACTGGCGCGCGCTGCTC
GGCCCGCTCGAGACCCTGCCTCTCCCTTTCTCC

(2) peptide sequence

Seq ID No 38 (>pEPOcos6_ORF2.pep)
MTSAVPTRQTSLLDDFERVADVDPERIAVHASETSLRYGDMNARANRIAHGLRARGIGPNQIV
AVAMARTPELMIVLYGILKAGAAYMPIARDAPPLRRDHMLRESQAALMIADEEIAGLAARVLT
PADPFFAAMPDHNPEPRHDPTDLIYVIYTSGSTGQPKGVAMEHRAVWNRLTWMQAQYPIDTQD
VILQKTPIVFDVSVWELFWWPLAGASVALLPQSMEKFPWAISATVARCGVTVMHFVPSMLMAF
LQVVAGRPEMADQMKGLRYVFCSGEALAPAHVSAFQEHINRAGSISLTNLYGPTEAAVDVSYF
DCPPGASLARVPIGRAITGIQLLVMRDGVPQPPGVEGELAIGGVGLARGYISRPDLTADRFVP
HPGGDGQRLYRTGDLVRRDADGELVFLGRIDHQVKIRGLRIEPGEIEAQISAHPDVADCALII
EQDSETLPKLTAYIVVARPGLTRKALLQFLGARLPDYMLPNRFLTLTELPVTANGKRDWRALL
GPLETLPLPFS

pEPOcos6_ORF3 sequences:

(1) nucleotide sequence

Seq ID No 39 (>pEPOcos6_ORF3.seq)

ATGTTACACCCGATTCCCACCGACCGTTTCGCCCTGAGCCGACCGCTCTTTCGCGGGTACCTC
GCGCACGATCCGATCGTGCAGGGCGTGCTGGCGGGCGACCATCCAGGCTGGGTCCTGGTGGAC
CGCGAGCCCGAGCCGCGCACGGCGCTGCTGTGGGCCTTTTCCGATCGGCTCTTCTGCGTGGGC
GCAGCTGACACGCTGACCCCGCACGCGCTGGCCGAGCTGTTCCACGACCGACTGATCCCCCAG
GCCCGTAAGATCGGGCAGCCGTTTTTCCAGGTTCAGGGCGAGACGGTCGACACCTGGTCGGAC
CACCTGCATCAGGTGTCGCCGCACGCGACAGTCTCCTTCCGCCAGGCATTCCGCTTCGACCGC
GACCTCTTCGAGCGGCTGCCAACCAAGCCGGAGCTGGCAGAGGCGCGGCTCGTGCCAATCGAC
GCGCGGCTGCTGGCCGAACAGGCTGATCTGCGCGAGCGGATACTGGCCTCCTGGTCCAGCGAA
GCTGCCTTCCATGCGCGCGGTTTCGGCTTCTGCTACCGCGTAGGTGACCAGCTGCCGAGCGTG
TGCCTGGCATCGCACGTAGGCGGCGGCGCGGCCGAGCTGAGCATCAACACCGAGCTCGAAGCG
CGCAATCGAGGTATGGCAACGCGGCTGTGCCGGCGTTTCATCGCCGAATCGCTGCAGCGCGGC
CTGACGCCTTGCTGGGGCACCGAGACCTTTCGCCTGCCGTCAATCGCGCTGGCCCAGAAGCTC
GGTTTCATCCCGACCTTCACCTTCCCCACCTACTGCTTCGCGACCGGCACCGAACAGCCGGAC
GACAACTTCCTAGGCGAGCTGTACTACAGGGAATCGCGCATCGCCGGAAGTGGGACCGATGAG
CCGCAAGCGGTTCGGCTGGCGCGGGGTTGGAGCCTGGCCGGCGACACCGAGCGTGCCGCGAGC
TTCGCCGCACGCGCCCTGGCCGAAGGGTGGGCCGGCCACTCGACTCTGGCCACCGATCCGGAT
TTCGCCCGATTGCGCGCCAGCGCCGCCTGGCCCCGCCTCAATGTCCCT

(2) peptide sequence

Seq ID No 40 (>pEPOcos6_ORF3.pep)

MLHPIPTDRFALSRPLFRGYLAHDPIVQGVLAGDHPGWVLVDREPEPRTALLWAFSDRLFCVG
AADTLTPHALAELFHDRLIPQARKIGQPFFQVQGETVDTWSDHLHQVSPHATVSFRQAFRFDR
DLFERLPTKPELAEARLVPIDARLLAEQADLRERILASWSSEAAFHARGFGFCYRVGDQLPSV
CLASHVGGGAAELSINTELEARNRGMATRLCRRFIAESLQRGLTPCWGTETFRLPSIALAQKL
GFIPTFTFPTYCFATGTEQPDDNFLGELYYRESRIAGSGTDEPQAVRLARGWSLAGDTERAAS
FAARALAEGWAGHSTLATDPDFARLRASAAWPRLNVP

pEPOcos6_ORF4 sequences:

(1) nucleotide sequence

Seq ID No 41 (>pEPOcos6_ORF4.seq)

ATGATTTGTCACTCCCACCGCTTCATTTTCCTCCACGTTCCCAAGGTCGCCGGCACAAGCGTC
AAGGACGTCCTCGGCCAAGAGCTATTCCAGGAGGACCAGGTCACGTTCCAGATCGCTCCCAAT
CCCCACTACCCACCTGAATGGACTGCGCCTTACGAGGAGCACATTATTGCCGCTGAATTGAAG
AGCCAGTTGGCGCCGGAAATTTGGGACGATTACTTCAAGTTCGCCTTCGTGCGCCATCCGCTC
GACTGGGCGGTCTCCAATTACTTCTTCTTCCTGCGCGACCGCAAAGGCCATCCGGCCCACGAA
TTCCTGGAGCGGAAGGGCTTCGCCGGTACCATGGACATGTTTTTCGGAGCGGCCGGGCGCCAT
CCGCTGGTCGCCGGCATGCGCTTCAGCCAATGGGAGTTCTTGTGCGACAGCGAGGGCCGGACG
CTGGTGGACTTCGTTGGCAAGTACGAGCGGCTCGAGCAGGACTTCGCCGCCGTGTGTATCCGC
ATCGGGCTGACCCCGCCCGACTTGCCGTGCCTCAACCAGACTCGCCACCAATCCTTTACCAGT
TACTACGACGAGGCTTTGATGCGCCAAGTCAGCCGCGCGTTAGCTCGCGATTTCGAAATTTTT
GATTATGCC

(2) peptide sequence

Seq ID No 42 (>pEPOcos6_ORF4.pep)

MICHSHRFIFLHVPKVAGTSVKDVLGQELFQEDQVTFQIAPNPHYPPEWTAPYEEHIIAAELK
SQLAPEIWDDYFKFAFVRHPLDWAVSNYFFFLRDRKGHPAHEFLERKGFAGTMDMFFGAAGRH
PLVAGMRFSQWEFLCDSEGRTLVDFVGKYERLEQDFAAVCIRIGLTPPDLPCLNQTRHQSFTS
YYDEALMRQVSRALARDFEIFDYA

pEPOcos6_ORF5 sequences:

(1) nucleotide sequence

Seq ID No 43 (>pEPOcos6_ORF5.seq)

ATGAAAGTGGACAAGCGGAATGTCGACGACATTCTCGGACTCACTCCGACACAGACAGGCATC
TTGTACCACTACCTGCTGGACCCGCAGGCCGACGCCTATTTCGAACAATTGACGCTGCACCTG

```
GAGGGGCCGCTCGACGTAGCGCGCTTCCGCCGCGCCTGGGAGCGCGTGGTGGCGGCTCACGAC
CAGCTGCGCGCCGTGTTTCGCTGGCAAGGGATCGAACACCCGGTGCAGATCATCCTCAAGCAG
CACGTGCCGGACCTGGAGTTGGCGGAGGTCCCGCGCGACGCCGATCCGGCAGCCTTCCTGGCG
CAATGGGTCGCGGCCGACCGGGCGCGCAAGTTCGACTTCGAGACGGTGCCCTTTCGCATCGGC
CTCTGCCGGACTGATACCCAACATCACGTGATGCTGCTCAGCAATCACCATATCCTGATGGAC
GGTTGGAGTACGGGCCTGATTCTGCGGGACTTCCTCGCCTGCTACGGCGACTCCGAAAACTGG
CGGCCACGCACCCGAACGCACTTCAAGGCGTTCATCAAGTGGCACCAGAACCGGCCACGCCGG
GGCGAGGAGCGATTTTGGCGCGACCTGTTGCGCGATGCGCCCGACGGCGGCTTTCCCCGCCTG
GGCGTCGAAGAAGGCACCCGCCACTCGCTTGACTTCGGCGCCCGCAGCCGCGCTCTCGACGAC
CGCTTGACCCAAGGCTTGCGCGACATGGCTCGCGACCTCGACGTCACCCTCGCCGCGATGCTC
CATACCGCTTGGGGCCTTCTACTCCAGCGCTACCAGAACAGCTGCGAAGTGATATTCGGGACC
ACCGTTTCCGGCCGCAACGTCGAGCTCGCCGGCCTCGACGAGGTGGTCGGCTTGTTCATCAAC
ACGATTCCGTTCCGCTTCTCGGCCGCGGCCGCGACGACGCCCGTCGAGGCCTTCCGTGCGGTA
CAGCGCAATCTGCTGGCGAGAAGCGAGTTCGAAGCCACCCCGCTGGTGGACATCAAGGGCTGG
AGTGGTCTCGGTCCGGGCGCGGAACTGTTCGACACCATCCTGGTCATCGAGAACTATCCCTTG
GACCGCGCTATCTTCGAGAGTGATTCCAGCCTGCGGTTGACCGACCACCAAATCTTCGAGCGC
ACCAATTACGGGCTGACCCTGACCATCGAGACCTTCAGCCGGTTGCACGTGACGCTAGCCCAT
CGCCGTGACCTGCTGGGCGACGCGGCCGCTGAGCGAATGCTAGATCATTTCACCGGCCTGCTC
CAAGCCATGCTGCGCTTCCCTCACCAGCCGTTCGCGCGCCTCGAGATGAAAAGCGAACACGAG
GCCCACCGCGTCCTGCACCAACTCAACCAAACGCGTCAGCCGCTGCCGTCCCAATCGGCTTTC
CACCAGTTGTTCTTCGAGCAGGCCCAGGCCGATGGGGCACGACCGGCGCTGTGGTGCGGCGCC
ACGCGCTGGACCTACGGCCAGCTGCTGGAACGTGCCCTGCGTCTGGCGGGACGGCTGCAGGAA
GCCGGCTTCGCCCGAGGCGATGTCGCCGCCGTCAGCCTCGGCCCGGTTCCGGATCTGATTCCC
GGTTTGCTGGGCCCGCTGTTCGCCGGCGGCGCCTACCTGCCGCTCGATCCCACCCTGCCGGCC
CAGCGCTCGCGGTTCATCCTCGACGATGCCGGTTGCCGCTTCCTGATCAGCGACGCGCCACTC
GCGGGGCCCACGCCGATCCATCCGGACCCTGCCGGCGCCAGCCCCGTTGACGTCATTTTTGCC
TGTCAGGACGGCGCCGCGCAGCCCGCCTACCTGATCTACACCTCGGGCTCCACCGGCCAGCCC
AAAGGCGTCTGGGTTAGCCACCGCAACCTGATCAACTTCCTGACGGGCATGAGCGCAATCCTG
CCGGTCGCGGCCGACGACGTGTTCCTCTCGCTGACTACCGTGTCGTTCGACATTTTCGGGCTC
GAGACGTGGTTCCCGCTCAGCCGCGGCTGCACGATCGTCTTGGGCACGCGCGCCGAGCAGTTG
```

GACCCGGCCGCGGCTGCCAAGGCCATCTCCTGCCATGGCGTCACGGTTTACCAGGCGACGCCA
TCGCGACTCCAACTTCAACTGGAGCACCCCACATTTGTCCGCGCCATCGGCTCCCTGACGACC
CTGCTGGTAGGCGGCGAACCCCTCCCAGCCGAGCTGCTGCGGCGCGTACGCGAAGTGACCGAT
GCGCGTATCTTCAACCTCTACGGTCCCACCGAAACCACCATCTGGTCCACAGCCGGGGAGGTC
ACCGCGGCGGACGTCCCGGATATCGGCCGCCCGATCGCAAATACCGGCGTTTTCCTTCTGGCG
CGAGACGGCTCGATCCAGCCGCCGGGCCTGGTGGGCGAGTTGTGCATCGCCGGCGAGGGCGTG
GCGTTGGGCTACCACCGACGGCCGGACCTGAACCGAGAACGGTTTCGCGAGATTCCGCCGGGC
CGCCTGCCCTTTGCCGGCAAGCTCTACCACACCGGCGACCTGGCCCGCTGGACCGAAGACGGA
CGGCTCCTCTGCCTGGGCCGTCTGGACGACCAGCTCAAAGTGCGCGGCCATCGCGTCGAGCCG
GGCGAGATCGAGGCAGTGATGGCGCGCCACCCGGCGGTCACGCAGGCGGTGGTCGTCACGCGG
CCGCGCAACGGCGAGCCGGTCTTGGTCGGGTTCTGGACTGCGGAAGGTGAGCCGATGCCAGAG
GAAGCGCTGAGCGCTTACCTGGCCGACCGACTGCCGAGCTACATGGTACCCGAACGGTGCATC
CTCATGAAGGCCATGCCGCTAACCGGCAACGGCAAGATCGACCGGCGCGCCCTACCCAATCCC
TTCGCCTTGACCGAGTCGACCCGGCAGGCGGCGCCGCGCACCTTGGCCCGCACCGCCGGCGAG
CATCGGGTTGCCGAGCTGTGGCAGGCCTTGTTGCGACGCGAGGCGATCGGCTTGGACGAACCC
TTTTTTCAGGCCGGCGGGAACTCATTCGGCTTGATTCGGCTTCACGCCAAGCTGGAATCCGCC
TTCGGGAAGTCGTTCCCGATCACCGATTTGTTCCAGCATACCAGTATTCGCAGCCAGGCAGAA
ATGCTGAGCGGCTCGTCCGTCGAGGCGCCGCTCGCGGGAGCCGTGCCGCAACCCCCGGCCGCC
GCCGCCCAAGTTGCCTCCTCGGCAGCTAAATCCCCAGGGGAGCGCGGCGCGGCAGCGACGTCG
AGCGGCCTGACCGCGCAACCGCCCCAACCCCACTTCCGGCCCATCGCCGTTATCGGCCTCGCC
GGCCGATTCCCCGCCGCACCCGACCTCGACGCCTTCCTTGAACTGCTCACGGAGGGTCGCTGC
GGCATTCGCTTCTTCAGCCAAGCCGAGCTGCGCGACGAGGGTCTCGACGCGAATCGAATCGCG
TGTCATAACTATGTCCCGGCCAAAGGTTTCCTCGACCGGGCCGACCACTTTGATGCCGACTTC
TTCGGCATCCCGCCGCGCGACGCAGAAATCACCGATCCGCAAATTCGGCTTCTGCTTGAGTGC
TGCTGGAACGCGCTGGAGCATGCCGGCTACCCGCCCGGCGGCGGCGAGATCGGGCTCTTCGCC
GGCTCCTCGGCCAACTATCACTGGCTCGAATACGTGGGCATTTCCGAGGAGAGCAGCAATCGA
TTCGCCGTCATGATTCAAAACGAAAAGGACTACCTGGCCACGCGGATCGCCTACCAGCTCGAT
TTGAAGGGCATTGCCGTCACCGTGCAAACGGCCTGCTCGTCGTCGCTGACCGCGGTCGAGCTG
GCCTGCGATGCGTTACACGCCGGCCGCGTGACCATGGCTTTGGCTGGTGGCGTTGGTCTGACC
TATCCGTTGCGCGCCGGATACCTGCACGAGGATGGAATGATCTTCTCCCCCGACGGTCGGTGC

CGGGCCTTCGACGCCCAGGCGGCCGGCACGGTCTGCGGCAACGGTCTGGGCATGGTGGTGCTG
AAACAGCTCGACGCGGCGCTGGCCGACGGCGATGCCATCCACGCTGTGATTAAGGGCATCGCG
GCCAACAACGACGGCGCGGCCAAGATCGGCTACACGGCGCCCTCGCAGAACGGTCAGGCGCGG
GTGATCCGCGCCGCCCATAGGCTCGCCCAAGTCGCGCCGGAGACCATCGGCTATGTAGAAGCC
CACGGTTCGGGCACGCCGCTGGGCGATCCGATCGAGGTGGCGGGCCTGACCGAGGCCTTTGAC
AGCCCGCGTCGCGGCTTCTGCGCCTTGGGTTCGGTCAAGTCGAATGTGGGTCATTTGGATGCG
GCAGCGGGCATCGCGGGTTTCATCAAGGCGGTGCTCTCGCTGTCCCATCGGACCCTGTTCGCC
AGCCTCCACGTCGACACGCCCAACCCGCAGATCCCGTTCGCCGACGGTCCGTTCCAGGTCAAC
ACGGAGACCCGGCCCTGGCCAGCTGCCGACCATCCCCGCCGCGCCGGCGTCAGCTCCTTCGGC
ATCGGCGGCACCAACGTGCACGCCGTCCTGGAAGAGGCGCCGCAGTTGGCCGAGCACGCGGGG
CGGCGGCGCGAGCGGCAGCTGTTCCTGGTCTCGGCGCGGACTGCAGCCGATCTGGAGCGACGC
ACCGCGGCGCTGGTCCGCCACCTGGCCGCGCATCCGGACCTCGCACCAGATGACGTTGCCTTT
ACCTTGCACGCGGGCCGCAAACCGATGACCCACCGTCGTTTCCTGGTCGCCGCCGACCTCGCG
GAAGCCGCCGCGCGTCTGGCCGAGCCCGATCCAGTCAAATCCGCCGCGGCGCGCGCCGACCGC
TGCCAGGTCTGGATGTTCGCCGGTCTCGGCTCTCAATACCCCGGCATGTGTGGCGGCCTCTAT
CGCACCGAGCCGGCCTTTCGCGAGCAAGTCGACCGCTGTTTCGACCTCCTCGCGCCGCGTTGC
GATTTGAAGCCCTCGCTCTTCCCCGAGCCCGATCAGGCCATCGACGCATCAGCCCTCGCGGCC
ATCGACACCGCCCAGATCGCCGTCTTCGTCTGCGAATACGCGCTCGCACGGATGCTGGAAGGC
TGGGGGCTGCGTCCGGATCGGCTGATCGGTTACAGTTTCGGCGAATACGTGGCCGCCTGCCTG
GCCGGCGTCTTCTCCCTGCCCGACGCCTTGGCAATCGTCCGCGAGCGTGGCCGGATCCTGGCG
GCGGCCGAGCCGGGCGCGATGGTCAGCGTGCCCCTTCCGGCCGAGCGCGTCGCGTCGCTGCTG
GAGCCGCCGCTTGCCTTGGCCATTGACAACGGCCCCTCATGCGTGGTGTCCGGGCCGGTCGAA
CCGGTGCGCACCTTCACCGCTCGCATGAAGCGGGACCGGGTCTGGGTGACGCCGCTCCAGGCC
GAGCGCCCGATGCATTCGCCGCTGATGGCCGAGGCCGGCGGCTCACTGCGCGCCATGTTGGCC
GGGTTCCGCCTGAATGCGCCGCGAATCCCGATCTTAAGCAATGTTACAGGAACCTACCTAACC
GACGAGCAGGCCCGAGACCCCGATTACTGGGCCCGTCACCTGTGCGGCAACGTTCGCTTCGCC
GACGGTGTGCGAACCTTGTTGGCCGAGCGCGATCCGGTGTTCCTTGAATTCGGGCCGGGCCGC
GATCTGAGCTCCTTGGTGCGCCACCAGATGCCGGAAGGCGCCGACGAGCCGATCGCACTGATC
CGTCATCGCGAAGATCCGGTGCGCGACGAAGACCTCCTGCTCGATGGCTTGGGCCGCTGCTTC
CTGCGTGGGGCGACCCTCCACGGGCAGGCCTTGTACGCCGGCCGAGGCTGCCGCCGCGTGCCG

```
CTGCCCGGTTACCCGTTCCAGGGTCCACGCTGCATGCCGGCCCGCGCCGGACTGCCCGGCCTG
GCGCGACCGACCGTGGGAGCGACCACCATCAGCTACCGACCAGCCTGGAAGCGGGCGCCGCGC
TTGGCGGCTGTCGAATCGCTCGCGCCGCAATCCTGGTTGGTATTCAGCGACGGCAGCGAATTG
GCGGGCGAGCTGGTGGCCGGCCTGCGCGCTTCCGGTTGCGCGACCACCCTCGTCGAAGGTGGG
CTGGCGTTCGCGCGCTTCGCGGGCGGCTTCCGCGCGAATCCCCGCGAGGAACAAGATCTCGCA
CAGCTGTTCGCGACCCTGTCGGCCGAAGCGATGCTGCCCACCCACATCCTGCACCTGCTCAGC
CTGCCGTCGCCGGAGCGCGACTCGCCGCTGGCGCGCCTGGAGCACCTCACCGAGCTGGGCTTC
CACCATCTGCTGGCCCTGGCCCGCCAACTGGAGGCGGTCGGCGCCCCCGAGGTCCGCCTCGCC
GTGGTGACAACCGGCCTGGCGGCGATTGGCGGCGAGTCCGAGCTGCGGCCCGAGGTCGGGCTG
TTGCGGGGACCTGTCCGCGTGATTCCCTTTGAATTCCCGAACTTGCGGCTGCGCCTGATCGAC
CTCGACTCGGCCGATCCCATCTGGCGTAGCGGTTGTGAGCCGTTGCTGCGCGAAATGGGCGCT
GCCCCGGGACCTGAAGAAATCGCGCTGCGCGGCACCAGCCGTTGGGAGTTGGGCTACGAGCCG
GTCGAGGGGGGCACCGTGAGCACCATCTCCTCGCGACTGCGCGAGGGCGGCGTCTATCTGATC
ACCGGTGGCCTCGGCGGCCTGGGTCTGGCCTTGGCCCGTCACCTCGCCCGGAAGTACCGCGCC
ACCCTGATCCTCGCTGGCCGGCGAGGCGCGCCGGCGCGCGAGCTCTGGCACCAGGCGCCAGCG
GAGTTCGTACCGGTCGCAGCTGCGATCGCACAGATGGAGGAGTGTGGCGCCCGCGTGATTCCC
GTCGCGCTCGACGTCACCGACGCCGACCAAGTGAACGCGTTGTTCGCCACCATAGAAGCTACG
GTCGGCAAGATTGAAGGCGTTTTCCACATGGCTGGCATCGTTGACGGCGGCATCATTCGAACG
CGCACGCGCGCTGCCAGCGACGCCGTGCTGGCGCCCAAAACGGTCGGAACCTGGATTCTCGAT
CGGGCTCTCCGCGGCGCCGGTGGCCGCTTCCTGGTGCTGTACTCCTCGATCAACGCGGTCGTC
GCGCCCTTCGGCCAGGTTGCCTACGCCGCCGCCAACGCCTTCCTCGACGCCTTCGCCAGCGCC
CACGAACACGACGAGCGTCTTTTCCGCGTCAGCATCGGTTGGGACACCTGGCGCGAGGCCGGC
ATGGCCGTCGATGCCGCCCGCGCCCGCGGCGACCAGGCCCCGCTCGAAGGGCTTAGCGACGAG
CAGGGCTTGCGCCTGCTCGAAAGCGCCTTGGTCGGTTGCGAACCGCGACTCCTCGTCTCCATC
AGCGAACTGCGCGCTCGACTAGCCGAGCATCATCGCAACGGCGGCATTCCCCGGTTGCTCGGG
CCCCGCGCCAACGAGGCGGGTGCAGCTGATTCCGGCGAGGAGGGCGCCACGCAAGACGCGTCG
CCGGCCCGTCGCGCCCGTCCCGATCTGGTCGTGGCCTTCGCGCCGGCCGGCAACGAGCTGGAG
CGCCGGATCGTGGCCATCATCGGCGCCTACCTGCGGCTCGGTCAGGTGGGCGTCGACGACAAC
TTCAACGATTTGGGCGCCACCTCGCTCGACCTCATCCAGATCGCCCAACGCCTCGGTCGCGAG
TTGGGCCGCGATGTCCCTGTCGTCTCGCTCTACCAACACCGCACCGTACGCGGGCTGAGCCGC
```

TTCCTCGGCGGCGCGCTCCAATCCGCGCGGTCCGGCGTCCCGACGGGCGCTGCCGCACCGGGC
GCCGCCACGCCGGGGGTTGCCACCCCGCCGCGGCCACAACCGTCGCGCCAGCACCTGGAAAAA
CGCCGTCAATTGAGGAAAAAAGGGGGGCCTTCCCATCATGAG

(2) peptide sequence

Seq ID No 44 (>pEPOcos6_ORF5.pep)

MKVDKRNVDDILGLTPTQTGILYHYLLDPQADAYFEQLTLHLEGPLDVARFRRAWERVVAAHD
QLRAVFRWQGIEHPVQIILKQHVPDLELAEVPRDADPAAFLAQWVAADRARKFDFETVPFRIG
LCRTDTQHHVMLLSNHHILMDGWSTGLILRDFLACYGDSENWRPRTRTHFKAFIKWHQNRPRR
GEERFWRDLLRDAPDGGFPRLGVEEGTRHSLDFGARSRALDDRLTQGLRDMARDLDVTLAAML
HTAWGLLLQRYQNSCEVIFGTTVSGRNVELAGLDEVVGLFINTIPFRFSAAAATTPVEAFRAV
QRNLLARSEFEATPLVDIKGWSGLGPGAELFDTILVIENYPLDRAIFESDSSLRLTDHQIFER
TNYGLTLTIETFSRLHVTLAHRRDLLGDAAAERMLDHFTGLLQAMLRFPHQPFARLEMKSEHE
AHRVLHQLNQTRQPLPSQSAFHQLFFEQAQADGARPALWCGATRWTYGQLLERALRLAGRLQE
AGFARGDVAAVSLGPVPDLIPGLLGPLFAGGAYLPLDPTLPAQRSRFILDDAGCRFLISDAPL
AGPTPIHPDPAGASPVDVIFACQDGAAQPAYLIYTSGSTGQPKGVWVSHRNLINFLTGMSAIL
PVAADDVFLSLTTVSFDIFGLETWFPLSRGCTIVLGTRAEQLDPAAAAKAISCHGVTVYQATP
SRLQLQLEHPTFVRAIGSLTTLLVGGEPLPAELLRRVREVTDARIFNLYGPTETTIWSTAGEV
TAADVPDIGRPIANTGVFLLARDGSIQPPGLVGELCIAGEGVALGYHRRPDLNRERFREIPPG
RLPFAGKLYHTGDLARWTEDGRLLCLGRLDDQLKVRGHRVEPGEIEAVMARHPAVTQAVVVTR
PRNGEPVLVGFWTAEGEPMPEEALSAYLADRLPSYMVPERCILMKAMPLTGNGKIDRRALPNP
FALTESTRQAAPRTLARTAGEHRVAELWQALLRREAIGLDEPFFQAGGNSFGLIRLHAKLESA
FGKSFPITDLFQHTSIRSQAEMLSGSSVEAPLAGAVPQPPAAAAQVASSAAKSPGERGAAATS
SGLTAQPPQPHFRPIAVIGLAGRFPAAPDLDAFLELLTEGRCGIRFFSQAELRDEGLDANRIA
CHNYVPAKGFLDRADHFDADFFGIPPRDAEITDPQIRLLLECCWNALEHAGYPPGGGEIGLFA
GSSANYHWLEYVGISEESSNRFAVMIQNEKDYLATRIAYQLDLKGIAVTVQTACSSSLTAVEL
ACDALHAGRVTMALAGGVGLTYPLRAGYLHEDGMIFSPDGRCRAFDAQAAGTVCGNGLGMVVL
KQLDAALADGDAIHAVIKGIAANNDGAAKIGYTAPSQNGQARVIRAAHRLAQVAPETIGYVEA
HGSGTPLGDPIEVAGLTEAFDSPRRGFCALGSVKSNVGHLDAAAGIAGFIKAVLSLSHRTLFA
SLHVDTPNPQIPFADGPFQVNTETRPWPAADHPRRAGVSSFGIGGTNVHAVLEEAPQLAEHAG

RRRERQLFLVSARTAADLERRTAALVRHLAAHPDLAPDDVAFTLHAGRKPMTHRRFLVAADLA
EAAARLAEPDPVKSAAARADRCQVWMFAGLGSQYPGMCGGLYRTEPAFREQVDRCFDLLAPRC
DLKPSLFPEPDQAIDASALAAIDTAQIAVFVCEYALARMLEGWGLRPDRLIGYSFGEYVAACL
AGVFSLPDALAIVRERGRILAAAEPGAMVSVPLPAERVASLLEPPLALAIDNGPSCVVSGPVE
PVRTFTARMKRDRVWVTPLQAERPMHSPLMAEAGGSLRAMLAGFRLNAPRIPILSNVTGTYLT
DEQARDPDYWARHLCGNVRFADGVRTLLAERDPVFLEFGPGRDLSSLVRHQMPEGADEPIALI
RHREDPVRDEDLLLDGLGRCFLRGATLHGQALYAGRGCRRVPLPGYPFQGPRCMPARAGLPGL
ARPTVGATTISYRPAWKRAPRLAAVESLAPQSWLVFSDGSELAGELVAGLRASGCATTLVEGG
LAFARFAGGFRANPREEQDLAQLFATLSAEAMLPTHILHLLSLPSPERDSPLARLEHLTELGF
HHLLALARQLEAVGAPEVRLAVVTTGLAAIGGESELRPEVGLLRGPVRVIPFEFPNLRLRLID
LDSADPIWRSGCEPLLREMGAAPGPEEIALRGTSRWELGYEPVEGGTVSTISSRLREGGVYLI
TGGLGGLGLALARHLARKYRATLILAGRRGAPARELWHQAPAEFVPVAAAIAQMEECGARVIP
VALDVTDADQVNALFATIEATVGKIEGVFHMAGIVDGGIIRTRTRAASDAVLAPKTVGTWILD
RALRGAGGRFLVLYSSINAVVAPFGQVAYAAANAFLDAFASAHEHDERLFRVSIGWDTWREAG
MAVDAARARGDQAPLEGLSDEQGLRLLESALVGCEPRLLVSISELRARLAEHHRNGGIPRLLG
PRANEAGAADSGEEGATQDASPARRARPDLVVAFAPAGNELERRIVAIIGAYLRLGQVGVDDN
FNDLGATSLDLIQIAQRLGRELGRDVPVVSLYQHRTVRGLSRFLGGALQSARSGVPTGAAAPG
AATPGVATPPRPQPSRQHLEKRRQLRKKGGPSHHE

pEPOcos6_ORF6 sequences:

(1) nucleotide sequence

Seq ID No 45 (>pEPOcos6_ORF6.seq)

ATGAGTGAAGTATCCATTCGCCCCGGCTTGGACATCGCGGTCATCGGCATGGCCTGCCGCTTT
CCCGGTGCCCGCAACCTCGCCGAGTATTGGGCCAACCTGATCGAAGGCCTCGAAACGCTCAGC
TTCTTCAGCGAAGAGGAGCTGCGTGAGGCCGGCTGCGATCCGGTCCAACTGGCCCAGCACAAC
TACGTGCGCACCAAGGGCCTGCTCCCTGACGCAGACCGTTTCGACGCCGATTTTTTGGTTAT
TCCCCGCGCGAAGCCCAGGTGATGGACCCCCAGATCCGCGTCTTCCACGAGGTCTGTTGGCAG
GCGCTGGAGCACGCGGGCTACAACCCGCATCGCCACACCGGCACGATCGGCCTGTTCGCCGGC
GCCGCGCCCAACGTTTTTTGGGAGTTTCTCTCCTATCGGTCCGATGCCGCCAATTTAGGCAAC

TTCACGCTGGGCCTGCACAACAACAAGGACTACCTGAGCTCGCGCATCGCCTACAACTTCAAC
CTGACAGGGCCCAGCTACACCCTGTTCACCGCCTGCTCGACCTCGATGGTCGCCATCCACCAG
GCCGTCCAGGCGCTGCTCAACGGCGAATGCGACCTGTGCATGGCCGGCTCGGTCTCCATTACG
CTGCCACTGGTTGCCGGCTACACCTACACGCCGGGCATGATCGTCTCGCCCGACGGCCATTGC
CGCACCTTCGACGCAGGCGCCAATGGCACTGTCTACGGCGACGGGGCCGGCGTGGTCGTTCTC
AAGCGGGCCGAGGATGCGTTGGCCGACGGCGACCACATATTTGCGCTCATCAAGGGCTCGGCG
CTCAACAACGATGGCAGTCGCAAGACCGGCTACACCGCGCCCAGCGTGCAGGGGCAGGTGGAG
GTGATCCGCGCGGCGATGAACCTGGCGGAGGTCGAGCCGGAGGCGATCAGCTACGTGGAAACC
CACGGGACGGGCACCACGGTGGGCGATCCGCTGGAGTTCGAGGCGCTAAAGGAGGCCTTCGGA
GGTGGCTGCAAGGCCTTCTGTGGATTGGGTTCGGTCAAGCCGAACATCGGCCATCTGGACGTG
ACGTCGGGGATCGCGAGCTTCATCAAGCTGGTCCTGGCGCTGGAGCACCGCATCCTACCGCCC
ACGCTCCACTTCCAACTGCCCAACCCGAAGATGGATGTGGTCGATAGCCCCTTCTACATCGTG
GCTGAGCGCGAACCCTGGCGCGAAGATCTGCTGCCGCGTCGGGCCGGTGTCAGCGCGTTCGGT
CTGGGTGGCACCAACGTCCACATGATTTTGGAGGAGTTTCAGCGCGAACCGGCGGCGAACAGC
GCGCGCACGCGCCACCTGACGGTGCTGACGGCGCGGTCGCCGCAAGCCCTGGCGCAGCTGGCG
GCCAACCTCGCCGAACACCTGCGCGAACACCCCGAGTTGGCGCTGGCCGATGTGGCCCATACG
CTGCTGCACGGCCGCAAGCCACATCCATTCGCGCGCATCCTGGTGGCGACCGATACGACGGCG
GCGATCGACGCCTTGATGAACGACCGCGATCCGCGAACGCGTTTCTTCGAAGCGACCGGGCGC
GGCGAGTCGGTGATCCTGTGTTTTGACGAAACGCCGCCGGAGCCGCGAAGCGCCCGCTACCTC
TGGGATCACGAGCCGCTTTATCGCGCGGCGGCGACGTCGTGCTTGGCTGGTGAGGTCGCCGAC
CCGGATCTGGAAGGCTGCTTTACTGCCCTGATCGCCGAGCAGGGCGCGGCAGCCGCCTTTTGC
CACCAATACGCGCTGGCCGGATGGCTGCTGGCCATGGGGTTGACCCCGTCGGCGTTGATCGGC
GTGGGCCAGGGCGAGTGGGTAGCAGCGGCGCTCGCGGAGGTGTTCCCGCCATCGGCCTGCTTG
CGCTGGATTAGGTTCGGCGAACGGCTCCCGCAGCCGCGCGATCAACGGATTCCGTTTCTCTCC
AATTTCTCTGGAAACTGGATCGTTGGGCGTGAGTTGGCCGACCCGGATTACCCCAGAAAGCAG
AAGGGTAAGCGCTGCATGAAGCGCCGTCGGTCCCAACCTCGGTCAGCTGGTGCAGGATGGGGG
CGATGGAACCGGCTCGGTCAGCTCGTCGCGCGCTGCTCTTCCGCGGGAAGCGGAGGCGGGACG
GTGATCGGCCCGAGGGCGAGGTTCATCTCGTCGTCGACGAGCCGGGCGCGGGTGCGCGCCCAG
TACCTGGGGGCGAGCTCGAGG

(2) peptide sequence

Seq ID No 46 (>pEPOcos6_ORF6.pep)

MACRFPGARNLAEYWANLIEGLETLSFFSEEELREAGCDPVQLAQHNYVRTKGLLPDADRFDA
DFFGYSPREAQVMDPQIRVFHEVCWQALEHAGYNPHRHTGTIGLFAGAAPNVFWEFLSYRSDA
ANLGNFTLGLHNNKDYLSSRIAYNFNLTGPSYTLFTACSTSMVAIHQAVQALLNGECDLCMAG
SVSITLPLVAGYTYTPGMIVSPDGHCRTFDAGANGTVYGDGAGVVVLKRAEDALADGDHIFAL
IKGSALNNDGSRKTGYTAPSVQGQVEVIRAAMNLAEVEPEAISYVETHGTGTTVGDPLEFEAL
KEAFGGGCKAFCGLGSVKPNIGHLDVTSGIASFIKLVLALEHRILPPTLHFQLPNPKMDVVDS
PFYIVAEREPWREDLLPRRAGVSAFGLGGTNVHMILEEFQREPAANSARTRHLTVLTARSPQA
LAQLAANLAEHLREHPELALADVAHTLLHGRKPHPFARILVATDTTAAIDALMNDRDPRTRFF
EATGRGESVILCFDETPPEPRSARYLWDHEPLYRAAATSCLAGEVADPDLEGCFTALIAEQGA
AAAFCHQYALAGWLLAMGLTPSALIGVGQGEWVAAALAEVFPPSACLRWIRFGERLPQPRDQR
IPFLSNFSGNWIVGRELADPDYPRKQKGKRCMKRRRSQPRSAGAGWGRWNRLGQLVARCSSAG
SGGGTVIGPRARFISSSTSRARVRAQYLGASSR

pEPOcos6_ORF7 sequences:

(1) nucleotide sequence

Seq ID No 47 (>pEPOcos6_ORF7.seq)

ATGGAACCGGCTCGGTCAGCTCGTCGCGCTGCTCTTCCGCGGGAAGCGGAGGCGGGACGGT
GATCGGCCCGAGGGCGAGGTTCATCTCGTCGTCGACGAGCCGGGCGCGGGTGCGCGCCCAGTA
CCTGGGGGCGAGCTCGAGGTAGCGGTCCCGCGGCCAGTAGGGCATCGCGCGAATGACGTCGGC
CAGGTAGGCCTCCGGGTCGAGCCCGTGCAGCTTGCAGCTCGCCACGAGCGAGAAGAGGTTGGC
CGCGGCGGAGGCGTGGTCGTCGCTGCCGAAGAAGAGCCAGGACTTTCTCGCAACCGCAATGGA
TCGCAGCGCTCGCTCGCTGGCGTTGTTCTCCAGGCGCAGCCGACCGTCGTCGAGGAAGCGCCG
CAACGGCTGCTCTTGGTTGAGGGCGTAGCCGAGCGCGGTGGAGACCAGGCCGCGCTCGCGGGG
ACGAGCGTGCTCGGCCCTGGCCCAGGCAAAGAACGCGTCGACCAGAGGGCGGACGACGACATC
GCGACGCACCTTGCGCTGCGCGGCGGCAGGTCCGCCAGCGCGCGATCGGCGGCAAAGAGGGC
GTTGATGCGCCGCAGCCCCTCGACACCGAGCTCGTGCTTGCAGACCGCCGCCTCCCAGAAGTT
GGTACGGCAATGCGACCAGCATCCGACTTCGGTCGGGGGCGGACCGCGCTTCTCGTCGGCAGC

AGCGCCTCTTGGTGGTGTGCCGCGGAAGAGGGCGTCATAGATGGCGTGAGCGTCAGCTTGAAT
ATACCGAGAGAAGCCGCGGAACATCTCGCAGACCGCGGCGCTGGTATGCTTGGGCTGGTACTC
GAAGAAGACGTGATCCTTGTCCGCGAGGACGACGAAGAAGTGTCCCTTGCGGCACGGCCCGGG
CTTCTTGTCCTTGCGCTCCTGGATGGGCCCAGGCTGGACGGAGACCCCGGTGGCGTCCGTGGA
CAGGCAGAAGGCGGTCTCGAAGGCCTCTTTGCGCGCGGCCTCGACGATGGCGCCCAGGGTCGC
ACCGACGTCTTCGGCGTAGCGGCACATCGTGCCGCGATCGAGCGACGCGCCCTGAAGCTCCAG
CTGCTGCTCCAGTCGATAGAACGGGACGCCGAGCAGGTACTTGCTGGTGAGGATGTGCGCAAT
CATCGACGGCGCGAGGAACGACCGCCGGAACAACTCCTTCGGAAGCGGCGTCGTGATGAAGAC
CGTGCAGGTCTCGCCCTTCGGCGCCGGCGGCGCGTCGAGCGAGGGCGCGTCGAGCGCTGT
GGAGGAAGCGCTCGGCTCGCCGGCCGCTGCCGTGTCCTCCGGGCTGACGCTCGCAGCGGGCGT
CGGCGTCGGGGCTTCTCTCGCGACGACCTGGAGCGGGGCCGCTTCCTCCTCGCCCGAACTGCT
CGCATCCGTGACGGACCGCTCGGCCTTGTACACGACGCGTGCGAGCACGATGCGGCGCATTCC
GCCGCGCTCGTAGCCGAGTCGCGAGGTCTCCTCGACCCCGATGCGCGTCGCCGTCGCATCGAG
CTCGGGGCAGGAGAGCTCGATGCGGACGACGGGCAGGTCGGACTCGGACAGGTCGCGACGGCC
CTTGCCGCCGGACCTTCGTTTCGGCCCCTTGGGGTCGTCGTGCTGCCGCTCGTCGCCTGTATT
GCGCTCGGCGGCGTCGAGTGCCTTCGCGAGGCGCTGGACCTCGAGGAACATCGAGTCGAACGC
CAGCTGCTCCGCGCTCACCTCGGCGCGCTCCGCCTTGGCCACGAACAGTCGACGTCGCAGAAG
CTGCAGCTGCTCGAGCGCACGGGTGTAGGCGCGCCGAAGCTGCGCGAGCGCATCGCGCGCTCC
CACGAGCTCGCTCTTTGCCGCGGCGAGCTCCGCTTCGAGCTGCGCGATGCGCTGCTGCTCGGC
CGAGAGCGTCGGCTTGGCGGCGGCGTCGTGCACGACGCCGCTCTACGTAAGCCGCGCGTACTT
GTCGAGCGAATTCGTGCGGCTCAGTGGACGCGGCGCGGTGCGCGCCTTCGCGGTTTGGACGTG
GGCGCGATCTCGATGCCGTCGAGCAGCGTCTCGAGCGTGGCGTCGTCCACCTCGACGTGCGTG
GCGCCCTCGGTCGGGGGGTCGGGAAGTGCGAACGCTCCGCGATCAAGGCGTTTTGAAAACAGG
CAGATTCCACTGCCATCGAAGAAGAGAATCTTGATCGTGGTCCGCCGCTTGCCGACGAACGCG
AACAGCGCTCCGCAGCGAGCCTCGTACCCCACACGCTCACGGATGAGACCCGAAAGCCGCTCG
AAGCCG

(2) peptide sequence

Seq ID No 48 (>pEPOcos6_ORF7.pep)

MEPARSARRALLFRGKRRRDGDRPEGEVHLVVDEPGAGARPVPGGELEVAVPRPVGHRANDVG
QVGLRVEPVQLAARHEREEVGRGGGVVVAAEEEPGLSRNRNGSQRSLAGVVLQAQPTVVEEAP
QRLLLVEGVAERGGDQAALAGTSVLGPGPGKERVDQRADDDIATHLALRGRQVRQRAIGGKEG
VDAPQPLDTELVLADRRLPEVGTAMRPASDFGRGRTALLVGSSASWWCAAEEGVIDGVSVSLN
IPREAAEHLADRGAGMLGLVLEEDVILVREDDEEVSLAARPGLLVLALLDGPRLDGDPGGVRG
QAEGGLEGLFARGLDDGAQGRTDVFGVAAHRAAIERRALKLQLLLQSIERDAEQVLAGEDVRN
HRRREERPPEQLLRKRRRDEDRAGLALRRRRRRVERGRVERCGGSARLAGRCRVLRADARSGR
RRRGFSRDDLERGRFLLARTARIRDGPLGLVHDACEHDAAHSAALVAESRGLLDPDARRRRIE
LGAGELDADDGQVGLGQVATALAAGPSFRPLGVVVLPLVACIALGGVECLREALDLEEHRVER
QLLRAHLGALRLGHEQSTSQKLQLLERTGVGAPKLRERIARSHELALCRGELRFELRDALLLG
RERRLGGGVVHDAALRKPRVLVERIRAAQWTRRGARLRGLDVGAISMPSSSVSSVASSTSTCV
APSVGGSGSANAPRSRRFENRQIPLPSKKRILIVVRRLPTNANSAPQRASYPTRSRMRPESRS
KP

pEPOcos6_ORF7.1 sequences:

(1) nucleotide sequence

Seq ID No 49 (>pEPOcos6_ORF7.1.seq)

ATGTTCCTCGAGGTCCAGCGCCTCGCGAAGGCACTCGACGCCGCCGAGCGCAATACAGGCGAC
GAGCGGCAGCACGACGACCCCAAGGGGCCGAAACGAAGGTCCGGCGGCAAGGGCCGTCGCGAC
CTGTCCGAGTCCGACCTGCCCGTCGTCCGCATCGAGCTCTCCTGCCCCGAGCTCGATGCGACG
GCGACGCGCATCGGGGTCGAGGAGACCTCGCGACTCGGCTACGAGCGCGGCGGAATGCGCCGC
ATCGTGCTCGCACGCGTCGTGTACAAGGCCGAGCGGTCCGTCACGGATGCGAGCAGTTCGGGC
GAGGAGGAAGCGGCCCCGCTCCAGGTCGTCGCGAGAGAAGCCCCGACGCCGACGCCCGCTGCG
AGCGTCAGCCCGGAGGACACGGCAGCGGCCGGCGAGCCGAGCGCTTCCTCCACAGCGCTCGAC
GCGCCCTCGCTCGACGCGCCGCCGCCGGCGCCGAAGGGCGAGACCTGCACGGTCTTCATCACG
ACGCCGCTTCCGAAGGAGTTGTTCCGGCGGTCGTTCCTCGCGCCGTCGATGATTGCGCACATC
CTCACCAGCAAGTACCTGCTCGGCGTCCCGTTCTATCGACTGGAGCAGCAGCTGGAGCTTCAG

GGCGCGTCGCTCGATCGCGGCACGATGTGCCGCTACGCCGAAGACGTCGGTGCGACCCTGGGC
GCCATCGTCGAGGCCGCGCGCAAAGAGGCCTTCGAGACCGCCTTCTGCCTGTCCACGGACGCC
ACCGGGGTCTCCGTCCAGCCTGGGCCCATCCAGGAGCGCAAGGACAAGAAGCCCGGGCCGTGC
CGCAAGGGACACTTCTTCGTCGTCCTCGCGGACAAGGATCACGTCTTCTTCGAGTACCAGCCC
AAGCATACCAGCGCCGCGGTCTGCGAGATGTTCCGCGGCTTCTCTCGGTATATTCAAGCTGAC
GCTCACGCCATCTATGACGCCCTCTTCCGCGGCACACCACCAAGAGGCGCTGCTGCCGACGAG
AAGCGCGGTCCGCCCCCGACCGAAGTCGGATGCTGGTCGCATTGCCGTACCAACTTCTGGGAG
GCGGCGGTCTGCAAGCACGAGCTCGGTGTCGAGGGGCTGCGGCGCATCAACGCCCTCTTTGCC
GCCGATCGCGCGCTGGCGGACCTGCCGCCCGCGCAGCGCAAGGTGCGTCGCGATGTCGTCGTC
CGCCCTCTGGTCGACGCGTTCTTTGCCTGGGCCAGGGCCGAGCACGCTCGTCCCCGCGAGCGC
GGCCTGGTCTCCACCGCGCTCGGCTACGCCCTCAACCAAGAGCAGCCGTTGCGGCGCTTCCTC
GACGACGGTCGGCTGCGCCTGGAGAACAACGCCAGCGAGCGAGCGCTGCGATCCATTGCGGTT
GCGAGAAAGTCCTGGCTCTTCTTCGGCAGCGACGACCACGCCTCCGCCGCGGCCAACCTCTTC
TCGCTCGTGGCGAGCTGCAAGCTGCACGGGCTCGACCCGGAGGCCTACCTGGCCGACGTCATT
CGCGCGATGCCCTACTGGCCGCGGGACCGCTACCTCGAGCTCGCCCCCAGGTACTGGGCGCGC
ACCCGCGCCCGGCTCGTCGACGACGAGATGAACCTCGCCCTCGGGCCGATCACCGTCCCGCCT
CCGCTTCCCGCGGAAGAGCAGCGCGCGACGAGC

(2) peptide sequence

Seq ID No 50 (>pEPOcos6_ORF7.1.pep)
MFLEVQRLAKALDAAERNTGDERQHDDPKGPKRRSGGKGRRDLSESDLPVVRIELSCPELDAT
ATRIGVEETSRLGYERGGMRRIVLARVVYKAERSVTDASSSGEEEAAPLQVVAREAPTPTPAA
SVSPEDTAAAGEPSASSTALDAPSLDAPPPAPKGETCTVFITTPLPKELFRRSFLAPSMIAHI
LTSKYLLGVPFYRLEQQLELQGASLDRGTMCRYAEDVGATLGAIVEAARKEAFETAFCLSTDA
TGVSVQPGPIQERKDKKPGPCRKGHFFVVLADKDHVFFEYQPKHTSAAVCEMFRGFSRYIQAD
AHAIYDALFRGTPPRGAAADEKRGPPPTEVGCWSHCRTNFWEAAVCKHELGVEGLRRINALFA
ADRALADLPPAQRKVRRDVVVRPLVDAFFAWARAEHARPRERGLVSTALGYALNQEQPLRRFL
DDGRLRLENNASERALRSIAVARKSWLFFGSDDHASAAANLFSLVASCKLHGLDPEAYLADVI
RAMPYWPRDRYLELAPRYWARTRARLVDDEMNLALGPITVPPPLPAEEQRATS

pEPOcos6_ORF7.2 sequences:

(1) nucleotide sequence

Seq ID No 51 (>pEPOcos6_ORF7.2.seq)

ATGATTCCGGCGGGCGTGCAGGTGTTCGTCGCGCTGGAGCCGGTGGACATGCGCTACGGCTTC
GAGCGGCTTTCGGGTCTCATCCGTGAGCGTGTGGGGTACGAGGCTCGCTGCGGAGCGCTGTTC
GCGTTCGTCGGCAAGCGGCGGACCACGATCAAGATTCTCTTCTTCGATGGCAGTGGAATCTGC
CTGTTTTCAAAACGCCTTGATCGCGGAGCGTTCGCACTTCCCGACCCCCCGACCGAGGGCGCC
ACGCACGTCGAGGTGGACGACGCCACGCTCGAGACGCTGCTCGACGGCATCGAGATCGCGCCC
ACGTCCAAACCGCGAAGGCGCGCACCGCGCCGCGTCCAC

(2) peptide sequence

Seq ID No 52 (>pEPOcos6_ORF7.2.pep)

MIPAGVQVFVALEPVDMRYGFERLSGLIRERVGYEARCGALFAFVGKRRTTIKILFFDGSGIC
LFSKRLDRGAFALPDPPTEGATHVEVDDATLETLLDGIEIAPTSKPRRRAPRRVH

pEPOcos6_ORF7.3 sequences:

(1) nucleotide sequence

Seq ID No 53 (>pEPOcos6_ORF7.3.seq)

ATGACAAGGACGAAGGCGACCGAAGTGATGTGGTCCGAGCGCGTTCGGGCGTGGCGCGAGAGT
GGTGAAACGGCGGAGGAGTTCGCTCGGAGCCGCGGATTTGCGGCCTCGACGCTGCACGGCTGG
TCGAGCCGGCTGTCGCGGGCCGAGCCACCGCGCTTTCTGCGCCTGGTGCCGAAGGCGCCCGCC
GTGACGAGCAGCGCTGCGGAGCTCGTCGTCGAGGTCGGCGGCGCGCGGGTGCGCGTCGCCGCG
GGGTTCGACCCCGCGCTGCTGGCGGAGGTGGTCCGTGCCCTCGGCGGAGCGGGGCGA

(2) peptide sequence

Seq ID No 54 (>pEPOcos6_ORF7.3.pep)

MTRTKATEVMWSERVRAWRESGETAEEFARSRGFAASTLHGWSSRLSRAEPPRFLRLVPKAPA
VTSSAAELVVEVGGARVRVAAGFDPALLAEVVRALGGAGR

pEPOcos6_ORF8 sequences:

(1) nucleotide sequence

Seq ID No 55 (>pEPOcos6_ORF8.seq)

ACTGGACAGCGCAGCCGGGGTGAGACGGCGCTTCGCGCAGCGCTTACGCAGAAGGCGCGCCGC
GCGCCATTGTCGGATGCGGTGCGCGACTTCGCCGCCGATCGGCTGTTGCTGGAACTGGGACAA
CCACTGGACGTAACGGCTGAAGCGAGCCAACGGCTCCAGCTCGCGCGGGGCGACCTGTTCGGC
GCCTACCAAGCGTTGGCCCAGCTCTGGATCTGCGGCGCCCTGGCCGAACCGCCGCGACTGTAT
CCCGACGAACACCGCCGGCGCGTGCCGCTGCCGAGCTACCCCTTCGAGGGAAAGCGGTTCTGG
ATCGAGGGCTCGCCGTTCGAAACCGCGCCCGCCGCCGGCGCCTCACCCCAACCCGCCGATTCG
GGGGACATTCTCAAGGGCGACCCGGCGGACTGGTACTATCGGCCGCGTTTCGAAGCGGCGCCG
CTCTTGCCCAGCCCGTTCGAGAGCGAACCCGGCGATTGGCTGGTGTTCGAAGATGAGCTGGGG
CTCGGCGCCTGGCTGAGCGAGACCTTGCGCGACAAGGGCGCGCGGGTCGCGACAGTCGTTCGA
GGCACCGAGTTCCGACGCCTGGCGTCACAGCGCTTCCAGCTTCGTCCCGATCGACGGGACGAT
TACCGGACCCTGCTGCACGAGTTGAAGGCGCAGGGCATCGCGCCGGTCCACCTGTGCCACCTA
TGGAGCGTGACCGCCGCACCGGATGCCGAGCAGTTGCTCGACGTCAGCTTTCACAGCCTGGTC
CATTTGGCGGCCGCTTTGGGTTCGGTTGGCTACTTCCACGCCATG

(2) peptide sequence

Seq ID No 56 (>pEPOcos6_ORF8.pep)

TGQRSRGETALRAALTQKARRAPLSDAVRDFAADRLLLELGQPLDVTAEASQRLQLARGDLFG
AYQALAQLWICGALAEPPRLYPDEHRRRVPLPSYPFEGKRFWIEGSPFETAPAAGASPQPADS
GDILKGDPADWYYRPRFEAAPLLPSPFESEPGDWLVFEDELGLGAWLSETLRDKGARVATVVR
GTEFRRLASQRFQLRPDRRDDYRTLLHELKAQGIAPVHLCHLWSVTAAPDAEQLLDVSFHSLV
HLAAALGSVGYFHAM

pEPOcos6_ORF9 sequences:

(1) nucleotide sequence

Seq ID No 57 (>pEPOcos6_ORF9.seq)

ATGAAGTTGAACGTGGTCGCCAACCGGCTATTCGACCCCGAGTCGCCCGAGCGCACCGAGCCC
GCCAAGAGTCTGTTGCTCGCGGTGACCAAAGTCCTGCCGCAAGAGGTGCCCAACGTTCGAACC
CGCGCCATCAGCGTGGACCTGGATCGCTCGTTCGACGCGGCGGCGCCCGCCTGGGCCGCCAGT
TTGTTGGTTGAATGCGGCGCGCCCGTCGAGGAAACGGTGGTGACCTACCATGGCGCAGCCCGA
TGGCTGCGCCGCTTCGATCGCGTTGCGGTGAATGGTCTCGGCCCGTTCCACCCCGATCAACCT
GCGCCGCTGCTGCGCGAGCGCGGCGTGTACCTGATCACCGGCGGCCTGGGCGGCGTGGCTGGC
CAGTTGGCGCGCTACCTGGCGCGGCCTGCCGGGCGCGGTTGGTGCTCACCGCGCGCCGGCCC
CTGCCCGAGCGCGACCAGTGGGATCGGGAGTCGGCCGTGCTGTCATGGGACGACAAGACGCGC
CAGCGCATCGAGCTGGTGCGCGAGCTGGAGCGGCTGGGGGCCGAAGTATTGGTGGTGGCTGCC
GATGTCGCCGACGAAGCGGCCATGGCGCAGGCGATCGAGGCCTCACTGGCGCGATTCGACGCT
TTGGACGGCTTGATCCACGGCGCCGGGATCGTGCGGGTCGCGTCGGGCCGCACGCCGATCGGG
AGTATGACGCGGGCCATGTGCGAGGAGCAGCTCCGCCCCAAGATGTTGGGCCTCGACGTCGTC
GACCGCCTCCTGCGCGATCGCCGGTTGGACTTCCGCATTGCCATCTCGTCGCTCGCCCCGATT
CTCGGCGGCCTCGGCCACGTCGCCTACGCCGCCGCCAACCTCTACATGGACGCGTTCGCGACG
CGCGCCGCCGCCGGCAACGCGCCTTGGATCGCGCTGAACCTGGCCGAGTGGGAATACGAGGGC
CCGGCTACCTACGACGAGCGGGTGGCCGTTCGCTCAAGCAGCTCGAGCTCACCAACGAGGAG
GGTATCCGCGTCTTCCAGACGGTGTTGGCCTTGGCCGCGCGCGGCCCGCTACAGCAGATCATT
ATTTCCACCGGCGACCTCCAGGCCCGCCTCGACAAATGGATTCACATCAAATCCCTGCATCGC
CGACCGGGGCCGGTCCAGCTCAGTCGCCGGACCGCGGCACCCCAGGGCGGTTTCGGCTCGGAG
CGCGCCGCCTTCGAGGCCGCCTTCGCTGACGCCTGGTGCGACTTCTTCGGGGTTGAAGAGGTC
GACCCGAACAAAAACTTCTTCGATCTGGGCGCCAGCTCGCTCGACTTCATCCACCTCGTCAGT
CGCTTCAGCAAGGCCATCGAACAGCATGTACCGCTCGAGGCCCTGCTCGAACACTCCACCCTG
CACGACCTCGCCGCCCACCTCGCGGGCGACGCGAACACCGACGCCAGCGACGAAGCGCGCATT
CGCCAACGGCTGCAAGGCGCCAAGTCCGGCGACATCGCCATCATCGGCATGGCCGGCCGCTTC
CCGCTCGCGCCCGACCTGGACACCTATTGGCGCAACCTGGTCGGAGGCATCGACGCGGTCAGC
TTCTTCAGCGCCGAGGAGTTGCGTGCTGCTGGCGTCACCGCGGCCGAGATCCACCACACCAAC

TACGTGCCGGCCAAGGGGCGCTGCGCCGACCAGGACTTGTTCGATGCGGCCTTCTTCGAATAC
ACTGCCAGCGACGCCGAGCTGATGGACCCGCAAAATCGCGTGTTACACGAGGTCGTGTGGCAC
GCGCTGGAAGACGCCTGTTTCGACTTCAACGGCGATCACGGCCAGGTCGGCCTGTTCGCGGGC
GCCTCGCCGAACCTGTGGTGGCAGTTCGTGGCCAGCTTTTCCGAGGCCGCCAAGACGCAGGGC
ATGTTCACCACCACCCTGCTCAACGACAAGGACTCGATCGCGACCCAGATTTCATACAAGCTC
GGTCTAAAGGGCCCCGCGGTCACCTTGTTCACCGGCTGTTCCACCTCGCTGGTAGCCGTTGAC
GCCGCCTGCCGCTCGATCTGGTCCGGTCAATCGGACATGGCCGTGGCCGGCGCGGTCTCGCTG
ACTCTCCCCGATAAGGCCGGCTACATCTACGAAAAGGGCATGCTCTTCTCGGCCGACGGCCAT
TGCCGGGCTTTCGACGCCAACGCCACCGGCATGGTCTTCGGCGACGGCGCCGGCGCGATCGTG
CTCAAGCCGTTGGACGCGGCCCTGCGCGACGGCGACCCGATCCATGCGGTGATCAAGGGCTGC
GCCACCAACAACGACGGCGACCGCAAAGCCGGCTACACGAGCGTCAGCGCCCAAGGCCAGGCC
GAGGTGATCCGCTCGGCCCAGATCCTGGCCGACGTGGCGCCCGAATCCATCAGCTACGTGGAA
GCCCACGGTACCGGCACCAAGTTGGGCGACTCGATCGAGATCAAGGCGTTGAAGCAAGCCTTC
GCCAGCGACAAGAACGGATTTTGCGGCATCGGGTCGGTCAAGACCAACCTCGGTCACCTGATG
GCGGCGGCGGGGATGGCCGGCCTGATCAAGACGGTTCTGGCGATGAAGCACCGCCAATTGCCG
CCATCGCTGCACTGCGACGAAGTGAACCCCGACCTGGAGTTGGAGCGCAGTCCGTTCTACATC
AACACCCGCCTGCGCGACTGGGTTGCACCGGGCGGGCCGCTGCGGGCCGGCGTGAGTTCGTTC
GGGATCGGCGGAACCAACGCTCACGTCATCCTGGAGGAGCCGCCGACGCGCGAGAGCGGCACG
CGCATGCGCCACTGGAAATTATTGATGCTGTCGGCGGCCAGCGAGGCGGCGCTCGACCGCCAG
GCCGATAACCTGGCCGACTACCTGGAGCGCCATCCCGAGGCCCACCTCAGCGACGTGGCCTAT
TCCCTCCAGACCGGCCGGCGCGTTCTGGCCTGGCGGCGCACGGTCCTATGCGAGTACCGCGAG
GACGCGGTGACCAGTCTGCGCGAGCGACAGGCCAAGCGCGTCCAGACAAGTCGCGTCCGCTGG
GACCACAAGGACGTGGTCTTCATGTTTCCCGGTCAGGGCGCCCAGTACCTCAACATGGGCCGC
GACTTATACGTCATGGAGCCGGTCTTCCGCGAGGTCATGGACCGCTGCTTCGAGTTGCTGGCC
CCTTTGTGGTCCGAGCATCCGCGCCAGATCCTTTATCCGGAGGGCGGGGTGTCGACCCTGCTC
CACCGGACTGATTACACCCAGCCGATCGTGTTCTGCTTCGAGTACGCCCTCGCCCATTTGCTG
CTCTCCTGGGGATTGAAGCCGGCCGCGACCATCGGCTACAGCTTCGGCGAGTACGTTTCTGCC
TGCCTCGCCGGCGTCTTCTCCCTGGAAGATGCGATCCGTCTGGTGACCGAGCGCGGTCGGCTG
ATGGCGGCTTTGCCCGCGGGCGCCATGCTCAGCGTCCCGGTTCCCGAATGCGAGCTGCTGCGG
CTGCTGGACGGCTTCCACGCCCAATCGGCGGCCCATCTGGCGCTGGCCGTCGACAATGGCGCC

TCCTGCATTGTGGCCGGCGAGCAGGCCGCCATCTCGGCCTTCGAATCGATGCTTCGCAAGAAG
CGTCTGTTGACCATGCGGGTCGCGGTCAGCCACGCCGCTCATTCGCAGGTCATGACCGGCGCG
ACCGACGCCCTGCGCAGCATCCTGCGGAAGATCCCCCTCTCCGCGCCGACAATTCCCTTCATT
TCCTGCGTCACCGGCACCTGGATCACTGCACAGCAGGCTACGGATCGCGAGTATTGGGTGAAC
CACATGTGCGGGACGGTGCGGTTCGCGGCGGGTCTGACCGAGCTGGGTCAAAACCGCGAGGCG
GTGTTCCTGGAAGTAGGTCCGGGCCGCGACTTGACGTTGCTGGCCCACCGCATCCTGGCCGAC
AGCGCGGCCGTGTTCGAGCTGGTCAAGGCGCCCGACGGCGGCGACGACGATGGGTTCCTCCTG
CTGGATCGATTGGCCAAGCTCTGGAGGCTGGGGATTTCGATTGACTGGGCCGGCTTCTACGCG
GATGAGCGGCGGCGGAAACTCTCGCTGCCGGGATATCCGTTCGAGCGGCGGCGCTTCTGGATC
GAGGGCAACCCGCTGGAGATCGCCGCCGGCAGGCCCAATGTCCAGGGGCCGCTGGTCAAGGCG
TCGGACATCGGCGCTTGGTTCTACGTGCCGCAATGGCGGCGGTCGGTGCTCGCCGAGCCGGGT
ACAACGGCGGCGGGCGCCGCCGTCACGGCGGAGCAGGCACGCGTCGTGACCGAGCTACGGGCG
GGATGCGCGTCGGCCGGCTTGGGCAGCGGGGCCTGCGGACTGAATGGCGGTGCCCCGTCCGAG
CGTCCGAAAGAAAGTGTAGCGCCAGCCGGGTCGACCAGCGCAGCGGCGCAGACCGGCGCGGAC
TGCCCGACACCGACTGGGGAGCCAGCGGCTGTGCCAAAGGACGGGGCCGAGCCGCGGCCGACC
TGGCTTATTTTCGCCGACGCCGGCGGATTGGCCGAATCTTTCGCCAAGCGGGTTCAGGCCCGC
GGCGAGAAGCTTTACCTGGTGGCTTCCGGCTCGCGCTTCGAGCGCCTGGCCGAGACCCGCTTC
CGCCTCGATCCCGGGGCCAAGTCCGATCACCGCCTGCTTTTCAAGGCGCTCGACGAGGCCGAC
ATCCTGCCGACCCACCTCCTCGACTTCCGCTCGCTTGACTGCGGCGGGCCCGACGCCGACCCC
ATGGACCAGGCCGGCTTCTTCGGGCTGTTGCACCTGGTCCAGGCGATGGCAGAGGCCGGCTAC
AGCCATCCCATTCGGCTGCTGATCGTCAGTTGCGGCGTCTACGATGTCACCGGTGCCGAACCG
CTGCAGCCGGCGCGGGCCACGATGATCGGACCGGCTCTGTGCATCCCGCAACAGTATCCGCAC
CTCGAAACGAGCCATGTGGATTTGGGCGTGGTCCATGCCGACGAGCTCCACGCCGCGCGCCAG
CTCGACAGCCTACTTGCCGAATGCCTAAGTGCAACGGCCGAGCGCCAATTGGCGCTGCGCGGC
CGACACCGCTGGCTGCTGGACTACGAGCCAGTCCGCTTGCCGCCGCTCGACCCGGGCCGTCTG
CCCTGGCGCCAGCGCGGGGTCTACTTGATCACCGGCGGTTTGGGCGGGATCGGCCGCATCCTG
GCCGAACACCTGGCCCGCACGACCTCGGCTCGCCTGGTCCTAATCGGCCGCGAAACCCTGCCC
GACCGCGACGACTGGGACGCCTGGCTGAACCGCCCGCAACCGGTCGACGCCACCCACGAACGG
CTGCTGCACAAGATCCGCGCGATTCGCGATCTGGAAGCGCTAGGCGCCGAAGTCCTGGTCCTC
GCCGCCGACGTCGCCAACGAAGCCGCCATGCGCGAGGCCTACGATCGCGCCGAATCCCACTTC

GGCACAATCCACGGGGTGATTCACGGCGCCGGCCTGATGGACGCGCAAAGCTTCTCACTGATC
GACGCCCTCGACCACGACCTCTGCGCCCGCCAGTTCGAAGCAAAAATCCGCGGCGTCTGCGTG
CTCGACCGCGTTCTGGCCGACCGCACGCTCGACTTCTGTCTGCTGATGTCTTCCATCTCCACC
GTGCTCGGCGGCCTGGGCTATTTCGGTTACGCCGCGGCCAACGCCTTCCTCGACGCCTTCGCC
CAGGCGCGCAGCCGCGACGCCGCTTTCCCCTGGCTTAGCGTGGCCTGGAGCGATTGGAAGTAC
TGGACCGAGCGCAAGATGGACAACGAGGTCGGCGCCGTCATCGACAGCCTCTCGATGGAACCC
GCCGAGGGCTTCGAAGCCGTCACCCGCGTCTTGGCTTGGGGCAAGGCGCCCCACATCGCCAAC
TCGCCCGGTGACCTCGGTCGCCGCCGGGATCAATGGGTCAAACTGGCCAGCCTGAAATCGGCG
CACTCCAGCGAGCCCGAGCCGGCTAGGCATGGACGTCCGGCGCTCTCCAGCGAATGGGTCGCG
CCGCGCAACGTGGTCGAAGAGAAGCTGGTCGCCATTTTCGAGCAGGTGTTCGGCACTGCGGCA
CTGGGCATCGAGGACAACTTCTTTGAGTTGCGCGGCGACTCGCTCAAGGCGGTCATGACCGCG
GCCCGTATTCAAAAGGAGCTGAACGTGGAAGTGCCGCTGCCGACCTTCTTCCAGATGCCCACG
GTCGCTGGCCTGGCCCAGTTCGTGACGCAAGCCAAGCGCAGCGGCCGGGAGACGATTCGGCGC
ACCGCGCCGCGCCCACATTACCCGCTCTCGGCTGCCCAGGGCCGCCATTACCTGCACTACCGC
ATGGACCCGCGTTGTACCGCATACAACGATCCCTTCGCCAACCTGATCGAGGGTCCGCTGGAC
GTGGATCGCGTGGAGCGCATCCTGCACACCCTCATCCTACGCCACGACTGCTTCCGCACCTCG
TTCCACTTCCGCGAGGGCGAGCCGGTCCAGGTGATTCACGATCGGGTGGACTTCAACCTGGCG
CGGATTACCTGCGCGCCCGAGGATTTGCCCGAACGGATGCGCGATTTCATCCGCTCCTTCGAT
CTGGAGCGACCGCCCGCCATGCGCGCCGGCCTCTTCGTCACGGGGCCCGAGCGCCACGTGCTG
CTAATCGATTTTCACCACATTATCACCGATGGCGTGTCGTTCGAGAACTTCGTCGGCGAGTTC
GCGGCGCTCTACCGCGGCGAGATCCTGCCCGAGCTGGAACTCGAGTACAAGGATTTCGCGGTG
TGGCAGCATGAGAACCGGGGCCGCCGCGCCAACAGCGACCAGGCCCGCTACTGGACCGAGCAG
TTGGCCAATGCGCCCGGGCCGATCGAGCTAACCACCGATTTCCCCCGTCCCAGTCGACGCAGC
TTCCGCGGCGACCGCGTGCGGACCGTGCTTGATGCGGAGCTCGTTGCTCGACTCAAAGAGCAC
GCGGCGCGCCTCGGCATCACCCTCTATAGCCTGCTGCTGGGCGGATTCTCGTTATTGCAGCAC
AAGCTCTCCGACTCGCACGACATCGTCATCGGTTCGCCCGTCGCGGGCCGCACCCGGAGCGAA
CTCCAGGATCTGCTGGGCGCGTTCGTCAACACCCTGCCGATGCGCCACCGCATCGACCCGACC
CATACCGCACGGGTCTTCTTGGAGCAGGTCCACCAGACAACCTTGGCGGCCCTCAGCTACCAG
GAGCACCCTTTTGACGAAATGGTGGCGACGCTCGGGTTCGCCGCCGATCCGGCTCGCAACCCG
ATCTTCGACACGATGTTCTTGCTGCAGAACATGGCCATGGGTGCAACCACCATTCCCGGTCTG

CGGCTCTCGCCTCACGACACTTTTCACCGCAAGGCATTGTGCGACCTGATGCTACAGGCGACC
GAGTATGACTGCCACCTGGAGCTGGTGCTCGAGTTCGCCACCGACCTGTTCCGGCTGGAAACC
GCGCAAGTCTTGCTCGACCGCTACCGCCAAGTCTTGGAGTGGCTGTTGGCGTACCCCCATGAA
TCGATAGACGATTTGACGCTCGCCGGCCACTTTCGCGAAGTCGAAGTGACGATGTCGGACGAG
GGCGACTTTGATTTCTCAGATTTCGAACCCCGCAACGTGAGAAACCTATGGCGCGCC

(2) peptide sequence

Seq ID No 58 (>pEPOcos6_ORF9.pep)

MKLNVVANRLFDPESPERTEPAKSLLLAVTKVLPQEVPNVRTRAISVDLDRSFDAAAPAWAAS
LLVECGAPVEETVVTYHGAARWLRRFDRVAVNGLGPFHPDQPAPLLRERGVYLITGGLGGVAG
QLARYLARACRARLVLTARRPLPERDQWDRESAVLSWDDKTRQRIELVRELERLGAEVLVVAA
DVADEAAMAQAIEASLARFDALDGLIHGAGIVRVASGRTPIGSMTRAMCEEQLRPKMLGLDVV
DRLLRDRRLDFRIAISSLAPILGGLGHVAYAAANLYMDAFATRAAAGNAPWIALNLAEWEYEG
PATYDERVGRSLKQLELTNEEGIRVFQTVLALAARGPLQQIIISTGDLQARLDKWIHIKSLHR
RPGPVQLSRRTAAPQGGFGSERAAFEAAFADAWCDFFGVEEVDPNKNFFDLGASSLDFIHLVS
RFSKAIEQHVPLEALLEHSTLHDLAAHLAGDANTDASDEARIRQRLQGAKSGDIAIIGMAGRF
PLAPDLDTYWRNLVGGIDAVSFFSAEELRAAGVTAAEIHHTNYVPAKGRCADQDLFDAAFFEY
TASDAELMDPQNRVLHEVVWHALEDACFDFNGDHGQVGLFAGASPNLWWQFVASFSEAAKTQG
MFTTTLLNDKDSIATQISYKLGLKGPAVTLFTGCSTSLVAVDAACRSIWSGQSDMAVAGAVSL
TLPDKAGYIYEKGMLFSADGHCRAFDANATGMVFGDGAGAIVLKPLDAALRDGDPIHAVIKGC
ATNNDGDRKAGYTSVSAQGQAEVIRSAQILADVAPESISYVEAHGTGTKLGDSIEIKALKQAF
ASDKNGFCGIGSVKTNLGHLMAAAGMAGLIKTVLAMKHRQLPPSLHCDEVNPDLELERSPFYI
NTRLRDWVAPGGPLRAGVSSFGIGGTNAHVILEEPPTRESGTRMRHWKLLMLSAASEAALDRQ
ADNLADYLERHPEAHLSDVAYSLQTGRRVLAWRRTVLCEYREDAVTSLRERQAKRVQTSRVRW
DHKDVVFMFPGQGAQYLNMGRDLYVMEPVFREVMDRCFELLAPLWSEHPRQILYPEGGVSTLL
HRTDYTQPIVFCFEYALAHLLLSWGLKPAATIGYSFGEYVSACLAGVFSLEDAIRLVTERGRL
MAALPAGAMLSVPVPECELLRLLDGFHAQSAAHLALAVDNGASCIVAGEQAAISAFESMLRKK
RLLTMRVAVSHAAHSQVMTGATDALRSILRKIPLSAPTIPFISCVTGTWITAQQATDREYWVN
HMCGTVRFAAGLTELGQNREAVFLEVGPGRDLTLLAHRILADSAAVFELVKAPDGGDDDGFLL
LDRLAKLWRLGISIDWAGFYADERRRKLSLPGYPFERRRFWIEGNPLEIAAGRPNVQGPLVKA

SDIGAWFYVPQWRRSVLAEPGTTAAGAAVTAEQARVVTELRAGCASAGLGSGACGLNGGAPSE
RPKESVAPAGSTSAAAQTGADCPTPTGEPAAVPKDGAEPRPTWLIFADAGGLAESFAKRVQAR
GEKLYLVASGSRFERLAETRFRLDPGAKSDHRLLFKALDEADILPTHLLDFRSLDCGGPDADP
MDQAGFFGLLHLVQAMAEAGYSHPIRLLIVSCGVYDVTGAEPLQPARATMIGPALCIPQQYPH
LETSHVDLGVVHADELHAARQLDSLLAECLSATAERQLALRGRHRWLLDYEPVRLPPLDPGRL
PWRQRGVYLITGGLGGIGRILAEHLARTTSARLVLIGRETLPDRDDWDAWLNRPQPVDATHER
LLHKIRAIRDLEALGAEVLVLAADVANEAAMREAYDRAESHFGTIHGVIHGAGLMDAQSFSLI
DALDHDLCARQFEAKIRGVCVLDRVLADRTLDFCLLMSSISTVLGGLGYFGYAAANAFLDAFA
QARSRDAAFPWLSVAWSDWKYWTERKMDNEVGAVIDSLSMEPAEGFEAVTRVLAWGKAPHIAN
SPGDLGRRRDQWVKLASLKSAHSSEPEPARHGRPALSSEWVAPRNVVEEKLVAIFEQVFGTAA
LGIEDNFFELRGDSLKAVMTAARIQKELNVEVPLPTFFQMPTVAGLAQFVTQAKRSGRETIRR
TAPRPHYPLSAAQGRHYLHYRMDPRCTAYNDPFANLIEGPLDVDRVERILHTLILRHDCFRTS
FHFREGEPVQVIHDRVDFNLARITCAPEDLPERMRDFIRSFDLERPPAMRAGLFVTGPERHVL
LIDFHHIITDGVSFENFVGEFAALYRGEILPELELEYKDFAVWQHENRGRRANSDQARYWTEQ
LANAPGPIELTTDFPRPSRRSFRGDRVRTVLDAELVARLKEHAARLGITLYSLLLGGFSLLQH
KLSDSHDIVIGSPVAGRTRSELQDLLGAFVNTLPMRHRIDPTHTARVFLEQVHQTTLAALSYQ
EHPFDEMVATLGFAADPARNPIFDTMFLLQNMAMGATTIPGLRLSPHDTFHRKALCDLMLQAT
EYDCHLELVLEFATDLFRLETAQVLLDRYRQVLEWLLAYPHESIDDLTLAGHFREVEVTMSDE
GDFDFSDFEPRNVRNLWRA

pEPOcos6_ORF10 sequences:

(1) nucleotide sequence

Seq ID No 59 (>pEPOcos6_ORF10.seq)

ATGGCGCGCCTGAGCCGCACAGATCTCCAACTCGCCATTCACCAGCGCACCGTGGAGCGCGAA
TATTGGCGCGCTCTGTTCGAGCGCCATCCGCAACGGTCCAGTTTGCCGGGGGTGCTCACCGCC
CCGATCGGCGACGAGTCGACCCGCGAGACCTTGTCATTCGTCCTCGACGAAGATCCCCTTCGG
CTGAGTAATCGTTCGCCGCAACGCCTGCTCACGGTGTTGGCGGCTGGCCTCGCGGCTTTCCTC
CACCGCTGCGACGGCGCTGAGCGCTTCACCCTGGGGTTGGCCCTACCGCGCCAAGCCGATGAC
CATCACCCGATCCTCAACAGCTTGATCGCGCTGGGGGTCGCGGTCGACTCGAGTACGACCTTC

CGCGATCTGCTCTATGCGCTTCGATCCGAATACCACGAGGCGATGCGCCACGCCAACTTTCCG
CTGGCGACCTGGTGGCGCGGCCTACCCGGCGGAACGGCGCCGTTCGACGTCGCCCTCAGCCTG
GACCCCTTCACAGACGGCGATTCGCTGGAAGACCACGCGATCGGCGCGTTGTTCCGGTTCGCA
TTGGAGGGTGAGCGCCTCACCTGCCGATTGCGATTCGACCCTGCGCGCTATGACCGTCCCGCG
ATCGAAAACCTCGCCGATCGTTTCGCCCGCTTCCTCACGCGCCTGTGCCGGGACGCCTCCACC
GTCATCCAGGCGCTGGACCTTTCGCTGCCAAGCGATGAATCGGTGTGGCGCGTCACTGAAGGC
GTGCGGCGCGGCTATTCGCAAGACCTGACGCTAGACCGCGCGTTCCGCCGCCAGGCCGCGCAA
ACGCCCGATCAGCCGGCGATCACGTTGAACGGGGACGTCCAGAGCTACGCCGAGGTCGACCGC
CGCAGCGACGCGCTGGCCCGCCACCTCCGTCGCCACGGCGTCGGTCCGGAAACGATTGTGGCC
GTCAACGCCCGGCGCGGGCCTAATCAGCTGACGGCCCTGCTCGCGGTCCATAAGGCCGGCGGC
GCCTACCTGCCGATCGATGCCGAGGAGCCGGCTGCCCGCCAGCAATTCAAGGTGCGCGACAGC
GGGGCGCGGTTGGCACTGGAGCCGTCGCCGGACCAGGCGCTGACCGTCACCGACCTGCCGCGG
CTCTTCCTGGACGATGCCTCGCTCTTCGCTGACGGCGGGCTCGATGTGCCGCGCGGCGCCGAC
TCGCTCAATCCGGCCTATGTGATGTACACGTCCGGCTCGACCGGACAGCCCAAGGGTGTGGTG
GTTCCCCACCGCGGCGTGGTCAATCGTTTGAATTGGGGGCAGTCCCGTTTCCCGCTGGACGAA
CGCGACCGAATCCTCCAAAAGACGCCGCTGCTGTTCGACGTGTCGGTCTACGAGCTGTTCTGG
GGCGCATGGAGCGGGGCCACCCTGGACATCCTCGAGCCCGGCGCCGAGCGCGACCCCGACGCA
GTGGCCAGGGCCCTGGCCGAGCGCGCCATTACCGTATGCCATTTCGTGCCTTCGATGCTGCTC
GTCTACTTGGAAGTCATGCGGCGGCACCATGCGCCGCCCGTGCCGGACCGGCTCCGTTACGTC
TTCGTCAGTGGCGAGGCCCTCGAACCGGACCACCTCGCCGGGCTCCAGCAGATTGGTCGGCGC
CTCGGCCGCACGATTCCCCTCGTTAATCTGTATGGACCAACCGAGGCCTCGATCGAAGTCTCC
TGCTTCGCCTGTCCCGCCGACCATGTGCCGCGCCGGATCCCCATCGGGCAGCCGATCGACAAC
GTCGCACTGCACGTTCTCGACCGGCGCGGCCGTCGCCAGCCGCCCTATCTTCCTGGCGAGCTG
TTCCTGGCCGGCGACTGCCTGGCGCGCGGCTACCTCAACCGTCCCGACCTGACCGCGCTCCAC
TTCGTGCCCAATCCCTTCGGCAACGGCGAGCGCATGTACCACAGCGGCGACTTGGCGCTCGTG
CGCGGCGACGGCCAAGTGGCGTTTCTCGGCCGCCGTGACCACCAAATCAAAATCCGTGGTCAA
CGGGTCGAACTGGGCGAAATCGAGAGTCATTTGCGCGGGCTCGAAGGCATCGCCGCCGCCGTC
GTCCAGGCCGAGTCGCAGCACCATGAAACCCTGCTGCACGCCTACGTCGTCACCAACGACGCG
GGCCTCAATGCGGCCCGGCTGCGCGCCGCCCTCGCTCAACATCTGCCCGAGTACATGATTCCC
CAGCGCTTCTCGCGGCTGGCCGAGTTGCCGCTGCTGGCGGCAGGCAAGATCGACCGCGCCGCC

CTCGCGCAACGTGCAACGCCGCTCGCCAGCGGCGCGCCCTTCGTGGAACCCAGCGGGCCCACC
CAGCAGCGTATCGCAGAACTGTGGCGCCAGGTCTTAGCGGTCGCCGAAGTCGGCGCCGAGGAT
CCCTTCTTCAGCATCGGCGGCAACTCGCTCAATGTGCTCAAGCTCAGCGCCGCGCTGAGCGAC
GCCTTCGCGCGTGACATTCCCATGCCGGCCCTGTTCCAATACGACACCATCGCCGCCCAGGCC
TCCTGGCTCGACGGGCAGGTTGACGAACGGGCCCAATCCGCCGCGCTCGACCGGCAGGCCGCC
GAGGCGGCGCTGACCCTTCAAGAGACCGTGGCCATTTTTGAGGGATTCGATGACGAACCA

(2) peptide sequence

Seq ID No 60 (>pEPOcos6_ORF10.pep)

MARLSRTDLQLAIHQRTVEREYWRALFERHPQRSSLPGVLTAPIGDESTRETLSFVLDEDPLR
LSNRSPQRLLTVLAAGLAAFLHRCDGAERFTLGLALPRQADDHHPILNSLIALGVAVDSSTTF
RDLLYALRSEYHEAMRHANFPLATWWRGLPGGTAPFDVALSLDPFTDGDSLEDHAIGALFRFA
LEGERLTCRLRFDPARYDRPAIENLADRFARFLTRLCRDASTVIQALDLSLPSDESVWRVTEG
VRRGYSQDLTLDRAFRRQAAQTPDQPAITLNGDVQSYAEVDRRSDALARHLRRHGVGPETIVA
VNARRGPNQLTALLAVHKAGGAYLPIDAEEPAARQQFKVRDSGARLALEPSPDQALTVTDLPR
LFLDDASLFADGGLDVPRGADSLNPAYVMYTSGSTGQPKGVVVPHRGVVNRLNWGQSRFPLDE
RDRILQKTPLLFDVSVYELFWGAWSGATLDILEPGAERDPDAVARALAERAITVCHFVPSMLL
VYLEVMRRHHAPPVPDRLRYVFVSGEALEPDHLAGLQQIGRRLGRTIPLVNLYGPTEASIEVS
CFACPADHVPRRIPIGQPIDNVALHVLDRRGRRQPPYLPGELFLAGDCLARGYLNRPDLTALH
FVPNPFGNGERMYHSGDLALVRGDGQVAFLGRRDHQIKIRGQRVELGEIESHLRGLEGIAAAV
VQAESQHHETLLHAYVVTNDAGLNAARLRAALAQHLPEYMIPQRFSRLAELPLLAAGKIDRAA
LAQRATPLASGAPFVEPSGPTQQRIAELWRQVLAVAEVGAEDPFFSIGGNSLNVLKLSAALSD
AFARDIPMPALFQYDTIAAQASWLDGQVDERAQSAALDRQAAEAALTLQETVAIFEGFDDEP

pEPOcos6_ORF11 sequences:

(1) nucleotide sequence

Seq ID No 61 (>pEPOcos6_ORF11.seq)

```
ATGACGAACCATGACCATCACGAGGAGAGCAGCGGCCTGGAGATCGCCGTCATCAGCATGGCC
TGCCGATTCCCGGGTGCTATGGCCTGCCGATTCCCGGGTGCTGCCGATTGCGACGCATTCTGG
GAAAACCTGATCAACGGGACCTCCTCGATCACCCATTTCAGCGACGACGAGCTGATCGCGGCC
GGCGTTGACGCGCGCGACCTGACGCCGCAGTACGTGCGCGCGGCCGGCCAGATCGATGACGCC
GAACGGTTCGACGCGGCCTTCTTTGGGTACTCCCAGCGTGAGGCCGAGCTGATGGACCCCCAG
TTCCGCCTGCTCCATGAATGCGCCTGGTCCTGTCTGGAACAGGCCGGCATCGATCCGCGCGTC
GAAGCCGCGCCGATCGGGCTGTATGCCGGCGCAGCCGACAACACCTACTGGAACGCGCTCTCG
TCGCTCGACCGGGGCTCGGCCGAATCGGAGCAATTCGCCGCCGAACAACTTTGCAACCGCGAT
TTTCTGTGCACGCTGGTCGCCGCCGCGCTCAACCTGAAAGGCCCCGCGGTGGTGGTTCAAAGC
GCCTGTTCGACCTCGCTGTTGGCGGTCCACTCGGCCTGTCGTGCGCTCCTGACCGGCGAATGC
CGAGTGGCCTTGGCCGGTGGGGTGGCGCTGCGCTTCCCACGCCCGAGCGGTTATCGCTACGAA
CCTGGCATGATCTTCTCGCCCGACGGGGTGTGCCGGCCGTTCGACGCGGGCGCTAACGGGACG
GTGCCCGGCGAAGGCGCGGGGCTGGTAGCGTTGAAGACGCTGAAACGTGCCCTCCAGGACGGC
GACACGATCCACGCCGTGATTCGCGCGACCGCGGCAAACAACGATGGTGCCCGCAAGACCGGG
TTCACCGCGCCCAGCGCCCACGGCCAAGCCGAAGTCATTCGCACGGCGCTGCGCCTGGCCCGG
GTGCCGGCCGAATCGATCGACTACGTCGAGGCCCACGGAACCGGCACGCCGCTAGGCGACCCG
ATCGAGGTAGCCGGCTTGGTGGAGGCCTTCGCCAGCGAGAAGCGCGGCTATTGCCGGCTGGGC
TCGGTCAAATCCAACCTTGGTCATCTGGACACTGCTGCCGGCATCGCCGGCCTGATCAAGACC
GTGCTGGCGCTCGAGCACGCGCACATCCCCAAGTCCTGCCACGTCGCCACGCCCAACCCCGCG
GCGCGCCTACACAAGACGCCTTTCCGCATTGCCGCCGACGGGATGGCCTGGCCGCGGCGTATG
GCGACGCCGCGGCGGGCGGCGGTGAGTTCGTTCGGCATCGGCGGCACCAACGTCCACGCGATT
TTGGAGGAGGCGCCGCCCCGCGCGCCCGAGCTGGCGGACGGGCGCAGTCAGGTGTTCGTCTTC
TCCGCCAAGGACGAGGCGGCGCTGGACCGTGCCCTTGCCAACTATGGTGCGGCCTTGGAGAAG
CGCGGCGACCTCGCGGCGGGCGCGGTGGCCTGGACGCTCCAAAACGGCCGGGCCGCATTCGAA
TGGCGAGCCAGCGCGGTGGCATCCGACCTCGACGAATTGGCGGGCGCATTGCGCGGCGAGCGG
CCCGGCGCCGTCAAGAAAAACCGAATGGCGCGCGAGGATAAGCCGGTGGCGTTCTTATGTTCG
```

GGGCAGGGGAGCCAGTACCGTGGCATGGGCCACGACCTGTACCGCGAAGAGCCGCGTTTCCGG
CACCACCTCGACGCCTGCCTCGCCATCCTCGCCGAACACAAGCCCGAGATCGACTGGCTGGCG
TTGCTGGGCTACCGCGACGAGGACGAGCCAACCGACCAGATCGGGACGTCCTCGCAGGGCCCG
AGCCGGTCAGCCGCATCGAACCCAGCGGAGCTCCTCGACAGCACCGAATTCGCCCAACCTTTG
CTTTTCTCCATGTCCTACGCGCTCGGTCGGCTGTGGCTCGACTGGGGCGTGCGACCCACGGCG
ATGATCGGGCACAGCCTGGGCGAGTACAGTGCTGCATGTATTGCAGATTTCTATGCACTCGAT
CAGGTGCTGCCCTTCATTCTGACCCGCGGTCGAGTCATGGCGCAATTGCGGCGCGGCTCGATG
TTGGCCGTCAGCGGTGACAGCGTTCTGATGCGCGAGCTGATCGCCGATGCGCTCGATTTGGCG
GCGATCAACGGCGCTGACCAATTTGTCTGGAGCGGGCCGAGCGAGGCTGTCCAAGCCGCGGGG
GTCCGACTGCGCGGCGCCGGCCTGCGTGCCACCGAGCTGAACACCTCACACGCGTTCCATTCA
GCCATGATGGATCCCATTCTGGAGGAGCTAACGGTTGCCGGTTCGCGACTTCAGGTCGGTGTC
GGGACGATTCCGGTCGTTTCATGCGTTACCGGAACCTGGTTGACGGCGAAGCAGCTGGCCGAT
CCGCGCTACCACGCGCGTCACGCGCGCGAACCGGTGCGGTTCGCGGCGGGCCTAGCGACGCTG
ACAGGGGAGGAGCCGCCGCTGATGCTCGAAGTGGGGCCGGGCTCGACCCTGGCGGCTTTGGCC
CGCGAGCATTCGAATGCCCGCCTCCCGGTCGTCACCAGCCTGCGCCACGCTCGCCAGGCGACG
CCCGATCGCCAATACCTGCTCGAAACGCTCGGCTGCCTTTGGCGACACGGGGTTTCCGTCGAT
TGGGGGGCCCATGCCGGACGTTCGCGACGCTTGGTTTCGCTGCCCGGCTATCCCTTTTCCGGC
GCGGTGCGCCGCTTAGCCGGCGACCCCCTCCGCCTGCTGGCCGGAGCCCGCGCCGTCGCCGCC
CCGTCGGGAACGCGCCAACTCAGCGCCGACGCGCGCGACCTCCCGAACACTCCGGAGCCGACA
TCCGGCGCCGTGTCGGCGATCAAAGCGCCAATCGCCGCCGCCGATCCCGGCCTCTATCGCCTC
TCCTGGCGCCAGGCCGGAACGGCGCCGCTCGGTCCGCCCGATCTCGGTCCGCCCCGCGACTGG
ATCGTCTTCGCCTCTGATTCTCACCTGCTCCAGGCGCTCAGGGCCAATCTCGGGACGCGCGCT
CAGCGGGTGACGCTGGTGACGCCGGGCCAGGAGTACGCAGCCGAGCCGTCCGGGTTTCGGCTG
CGGCCGGACCAGATCGACGATTACCGCGCCCTGTGGGCGGACTTGGCGCAAACCGGTATTGTG
CCACGATACATCGCGTTCCTCGCCCCGTTCATGTACCGGGCGCGCATGGCGGGCGATGCCTCG
ACCCTGGACGAAGTGCGCGAGGGCGGCTTCCTGCCCCTGACCCGCTTGATCCAGACTCGCCCG
CCAGGCGGACCGAGCGGACTTCTAAGCCTCACGATCGTCACCCCGGCCGCCCTGGCGCTGGGC
GACGAAGCGACGCGCCCGGAATGGGCAATCCTGCACGGGATGGTCGCCGGCTTAAGCCGCGAT
TATCCCGAATGGCGCTTCGTCTCGATCGACGGCGGCGACCCATCCCCGCATCGGTGCGAAGGT
CTGGCCCGCTTGATCGCGCTTCATGCGGTCGACGAGGCTGGCCCGACCCGCTTGGCGCTGCGC

GGCCTTCACGCTTGGGTTCCACAGTGCGAGCACGTTCAGCCGGCCACCATCCCTGGGGCGGGT
ATGTGGCGCGAGGGTGGTGTGTACATGATAACGGGCGGATTCGGCGGGATCGGTCTGGCGCTG
GCCCGCGCCCTGGCTCGAGAAGCTCGCGCCAAGCTGATCCTGGTCGGCCGAAACCTGCCCACC
GCGCCGATCGATCTCGAGGCTTGGACGCGCCGCCGTTGATTCTCACCGCCGACGTCGCCGAC
GAAGAGGCCATGCGCCGCGTCTTCGATGCCGCGCACGCCCGGTTCGGCGCCATCGACGGCATT
CTTCACGCGGCCGGTGTCCCCGGTGGCAGCCTGTTCGCCAACCAATCGGACGCGGCCTTCGAA
GACGTGCTGCACGCCAAGGTTCGCGGTACCCTCGTGCTGCAAGGCCTGAGGGCAATCGATGCG
CCGCTGTTGCTGATGTCCTCGCTGGACGCCTGGCTTCCCGGTCCCGGTCAGACCGCCTATGCC
GCCGCCAACGCCTTCCTCGACGCCTTCGCCAGTCTGCGCCGGCGAGAGGGAGAGCCGGTGTAC
AGCGTTGGCTGGGACAGTTGGTGCGAGGTGGGCATGGCTGCTCGGGTCGCTGCCCGATCGGCC
GACGAACGCGGCCGCCTGGCGCGCGAGGGGATCAGCCCTCGCCAGGGTTGGCAGGCTTTGAGC
CGGGCGCTCGCCCTCGACCCCCCCCACCTGATGATCTCGCGCACCGACCTGACCTCGCGCTGG
CACAGTCGATCCAGCCCTACGCCGGTCGCCTCGAGCGAACCCGAGGTGGCGCTGCCGCGCTGG
ACCGCATCCGCCTGCCAAGCCGTCATCGAGCGTGTTTGGTGCGAGCACTTCGCCACCGCCGCC
GTGCCTCCCGATGGCAACTTTTTCGAGCTCGGCGCCAGTTCCTTCGACATCGTCCAGCTCAGC
GCTCGACTTCAACAACAGTTCGGCCGAGATGTCAGCCACACCGTGCTCTACAGTCATCCCACC
GTCGCCTTGCTGGCCGGCTACTTCGCCAATGACCCGACGCCGTCCGGTGCTGCTGCCGACGAA
CGCGACGAAGCGGTGCGTCGCGGCCGCGACCTCTTGAAGAGCCGCCGGCGAGGAGTA

(2) **peptide sequence**

Seq ID No 62 (>pEPOcos6_ORF11.pep)

MTNHDHHEESSGLEIAVISMACRFPGAADCDAFWENLINGTSSITHFSDDELIAAGVDARDLT
PQYVRAAGQIDDAERFDAAFFGYSQREAELMDPQFRLLHECAWSCLEQAGIDPRVEAAPIGLY
AGAADNTYWNALSSLDRGSAESEQFAAEQLCNRDFLCTLVAAALNLKGPAVVVQSACSTSLLA
VHSACRALLTGECRVALAGGVALRFPRPSGYRYEPGMIFSPDGVCRPFDAGANGTVPGEGAGL
VALKTLKRALQDGDTIHAVIRATAANNDGARKTGFTAPSAHGQAEVIRTALRLARVPAESIDY
VEAHGTGTPLGDPIEVAGLVEAFASEKRGYCRLGSVKSNLGHLDTAAGIAGLIKTVLALEHAH
IPKSCHVATPNPAARLHKTPFRIAADGMAWPRRMATPRRAAVSSFGIGGTNVHAILEEAPPRA
PELADGRSQVFVFSAKDEAALDRALANYGAALEKRGDLAAGAVAWTLQNGRAAFEWRASAVAS
DLDELAGALRGERPGAVKKNRMAREDKPVAFLCSGQGSQYRGMGHDLYREEPRFRHHLDACLA

ILAEHKPEIDWLALLGYRDEDEPTDQIGTSSQGPSRSAASNPAELLDSTEFAQPLLFSMSYAL
GRLWLDWGVRPTAMIGHSLGEYSAACIADFYALDQVLPFILTRGRVMAQLRRGSMLAVSGDSV
LMRELIADALDLAAINGADQFVWSGPSEAVQAAGVRLRGAGLRATELNTSHAFHSAMMDPILE
ELTVAGSRLQVGVGTIPVVSCVTGTWLTAKQLADPRYHARHAREPVRFAAGLATLTGEEPPLM
LEVGPGSTLAALAREHSNARLPVVTSLRHARQATPDRQYLLETLGCLWRHGVSVDWGAHAGRS
RRLVSLPGYPFSGAVRRLAGDPLRLLAGARAVAAPSGTRQLSADARDLPNTPEPTSGAVSAIK
APIAAADPGLYRLSWRQAGTAPLGPPDLGPPRDWIVFASDSHLLQALRANLGTRAQRVTLVTP
GQEYAAEPSGFRLRPDQIDDYRALWADLAQTGIVPRYIAFLAPFMYRARMAGDASTLDEVREG
GFLPLTRLIQTRPPGGPSGLLSLTIVTPAALALGDEATRPEWAILHGMVAGLSRDYPEWRFVS
IDGGDPSPHRCEGLARLIALHAVDEAGPTRLALRGLHAWVPQCEHVQPATIPGAGMWREGGVY
MITGGFGGIGLALARALAREARAKLILVGRNLPTAPIDLEAWDAPPLILTADVADEEAMRRVF
DAAHARFGAIDGILHAAGVPGGSLFANQSDAAFEDVLHAKVRGTLVLQGLRAIDAPLLLMSSL
DAWLPGPGQTAYAAANAFLDAFASLRRREGEPVYSVGWDSWCEVGMAARVAARSADERGRLAR
EGISPRQGWQALSRALALDPPHLMISRTDLTSRWHSRSSPTPVASSEPEVALPRWTASACQAV
IERVWCEHFATAAVPPDGNFFELGASSFDIVQLSARLQQQFGRDVSHTVLYSHPTVALLAGYF
ANDPTPSGAAADERDEAVRRGRDLLKSRRRGV

pEPOcos6_ORF12 sequences:

(1) nucleotide sequence

Seq ID No 63 (>pEPOcos6_ORF12.seq)

ATGACCGTGGAGCACGAAACCGGATTCGAAATCGCCGTCATCGGGCTGGCTTGCCGCGTTCCC
GGCGCTGCCGACGTGGCCGCCTTCTGGCGCAACCTGGTCGAGGCCAAGGAGAGCGTGCGCTTC
TTCGAGGACCACGAGCTGCGGGCCGCCGGCGTGCCCGAGGAGATCTTGCGCCTGCCCAACTAC
GTGAAGGCCAAGCCACTGCTCGCTGATGGCGAAGCTTTCGACGCGGACTTCTTCGGGTTCCAT
CCGCGCGAGGCCGCCTACCTGGACCCGCAAGTTCGGCTCCTGCACGAATGTTGTTGGACCGCG
CTGGAGGATGCCGGCTACGATCCCGCGCAGTACGCCTACCCGATCGGGTTGTTCGCGGGCGTC
TCCAGCAATCTCTCGTTCCTGTTCGACCGCATCGATCCGCGCGACTCCCCCCTGCAGAAGCGC
TATGTGGCCGAGCTGAACGCGGCCTCCTTCGCCACCCAGATCGCCTACCGGCTCGATCTGAAG
GGGCCGGCCATTTCGATTCAAACCGCCTGTTCGACGTCACTGGTGGCGATTCACCTGGCGGCG

CAAAGCCTGATCGGCGGCGAGTGCCACATGGCCTTGGCCGGCGGAGCGACCTTGGAGGTCCCC
AAAAAGCCCGGCTATCTCTACCGCGAAGGCTACATCAACTCGCCGGACGGCCACTGCCGGGCC
TTCGACGCCGACGCGGCCGGCACCATCTTCGGCGACGGCGTCGGCATCGTCCTGCTCAAACGC
TACCGCGACGCCCTACGCGACGGCGATCACGTGTACGCAGTGATCAAAGGCTCGGCGATCAAC
AGTGACGGCCATCGCAAGGTGTCCTACACGGCGCCGGGCAAGAGCGGTCAAGTGGCGGTGATC
CGCGCTGCGCTGGCGGCGGCCCAGGTAGAGCCGCAAACCATTCGCTTCGTCGAGGCCCACGGG
ACCGGCACACTCGCCGGCGATCCGATCGAGGTAGAGGCGTTGACGGAGGTCTTTGCCGAAGCG
GGTCGCGGTACCTGCGCCCTGGGTTCGGTGAAGACCAACATCGGCCACTTGGATGTGGCGGCG
GGCGTGGCCGGTTTCATCAAGGCGGTCTTGGCGCTCGAGCGGCGCGTCCTCCCGCCCAGCCTT
CACTTCGTCCGGCCCAACCCGGCCATCGATTTCAACGGGCCCTTCTACGTTTGTCGCCAAATC
GAGCGGTTGACGGAGAACGGGCGGTTGCGGGCCGGGGTGAGTTCCTTTGGCATTGGCGGCACC
AATGCCCACGTGATTCTGGAGGAAGCGCCGGCGCCGGAGGCGAGACTGCCGGCCGGGAGCCCG
CCAGGCGCGAGTCCGTTCCTGTTCCCGCTATCGGCCAAGACGCCGGATGCGCTGGCAGGCCGT
TGCCACGACCTTGCCGACCACCTGCGGGCGCACCCCGAGCTCCTCCTGGCCGATGTGGCCCTC
ACTCTGCAGATGGGGCGGGCGTCGTTCGCCTACCGCCATGTGGTCCAGGCTGCGACGGCGGAG
GAGCTGATTCGCGGTCTGGGAGCGTTCCGACAGGAGTCCATCCGCAAGAGGCGGAATCGAGTA
CAATGGGTGTTGGCAGGCGAGGCGATGTCGCTTGACGCCGGTTTGCGGCTGTACGCCGATTGG
CCGGTCTATCGGGAGCGGGTCGACGTCTGTCTGGCGATCGTCGCCAAGCTGCGCCAAATCGAC
GGCCGGTCATTCCTACATGAGTGGATCGAGCGACCGCGCGAGGTTCCTGCCGAATGGTCGACG
GCGCTGGCGTTCATGTTCCACTGCGCGCTGGCGCAAGCCCTGAGCCAGGCCGGCCTGCACCCG
CAGCGCATGTGGAGCCGTGGGCTGGGCGGACAGGTCGGCGTGGTTTTGGCCGAATCCCTGTCG
TTGGAACAAGCGCTGGCGCTGGTGTTGTGCCAGACACCGGTTCCCGGCGATGCCACACCTCAG
CGCGAACGCTTGGTTCGGACACTGGAAGGCTGCCGGTTTCGTCCACCACGATTTTTGATTTCG
GCAGACAGCTCGGGTCGACCCCTGGACCTCGCCGAATTCGCTCATGTCGATTTTTGGTGCGGT
GGCCAAAGCGCCTCGCCCAATGAGGCGGAGCTGCGCTCATGGAGCGACGCCGCGCCCGAGCTG
GTGACCTTGGCGATCGGCCCATCCTTTCTCGAGGCCGCCTCCGGGACGGTGGGTCTGGCGATC
GACCCCAAGCGACCGATGACCTGTGTTCAGCGCACGGTGGCCGCGTTGTGGAATGGGGATGT
GACGTGCGCTGGGCTGCGTTCACCTCGTCGACCGGGCGTCGGGTTCCCCTGCCTACCTATCCC
TTCGTGCGGGTAATTCCCACGATCGGCGACCCCCTTCGCGGAGCAGGCGCGGAGGATGACTTG
ATTGCGGCGAGCGCTTCCGCGTCGGCCGGATCGCCGCCCGAGCCGTCGGCAAACTCGGCAGCG

GAACGCCCACGCGCCCAGTCAAGCATCGCCTCGGCAACCACACCGGCTCCGTCTCATACGTCG
GCCAGCGTGGCCGTGGCCACCATTCTCGAAACCGTCCGTGCCTATTTCGGGTTCGCCGCCGTG
CGTTCCACCGACGCCTTCTTCGAATTGGGCGCGTCCTCGCTGGATTTGGTCAACCTGGGCCAG
CTCCTTTCCGATCGTCTCGGCCGCGAGGTTCCGACCCTGCTCCTCTACGACCACCCAACACCG
GACCAGTTGGCGCTGGCCCTGACATCCGCGGCGCTCAGCGCAGAGGCGCCGCCCTTAAGGGGC
GGTCATCGCGCATCGACTTCCGGCACAGCCGCGAGCTCGGCCGCCTCCACCGCACCGACGTTC
CCGGGGGACGCTCACTCGCAGCCCAGCTTCGTTCGCGAGCAGGACATCGCCATCATCGGGATG
GCCTTCCGGGGACCGGGCGCCGACGACCTGGACGCGTTCTGGAACAACCTGGTCGAAGGGGTC
GAGTCGATCACCTTCTTCAGCGAGGACGAGCTGCTGGCGGCGGGCGTCCCCCGCGAACATCTG
GCCTCGACGCGCTACGTGCGGGCCAAGGGGGAACTGACTGGGATGATGGATTTCGAACCGGAA
TTTTTCGGTTATTCGGCGCGCGAGGCGGCGGTCATGGACCCGCAGTTCCGCGTGTTCCACGAA
TGCTCCTGGCACGCACTGGAGCACGGCGGCTACGATCCGACCCGATGCGCGGCATCGATTGGC
GTCTACGCCGGCGTGACCAACCACCTGCCTTGGCTGATGCGAACTTTGCCGCACCTGACCGAG
GAGGAGCAATTCGGCGCGCTGCTCCTCACCGACCGCGAGTTTTTCGCACCGCTGCTCTCCTAC
AAGGTCGGCCTGCGCGGACCCGCTATTTCGCTGCAAACCGCCTGTTCGACGTCGTTGGTGGCG
ATCGGCACGGCCTGTCGCGAATTGCGCGCGGGTGCCTGTCAGATGGCCCTAGCGGGCGGCGTG
ACGGCCAGCATCGAGCGCTGCGGCTACTTCCACCAAGAAGGCTACATCCTCTCGCCTGACGGC
CACACGCGCAGCTTCGACGCGGCGGCCGCCGGCACGGTCTTCGGCGACGGAGTCGGCATGGTG
CTGCTGAAGCCGCTGGCCCAAGCCTTGGCCGACGGCGACACGATCCACGCGGTGATCAAGGGA
ATCGGCATCAACAACGACGGCGCGCGCAAGGTCGGCTTCACCGCACCTAGCCGGGCCGGTCAG
ACCGAGGCGATTCGGGCCGCGCTGCGCGACGCCGGGGTGGCGTCGAACCGCGTCAGCTACGTG
GAGGCGCATGGAACCGCGACCAGAATGGGCGACCCGATCGAGGTCGAGGCCTTGACCCAAGCC
TTTCGCGCCGAAGCCGACGGTCCGCTTCCGCCCGGCTCCTGCCTACTCGGCTCGGTGAAGTCC
AACGTGGGCCACCTGAACGCCGCGGCCGGCGTGGCTGGTCTGGTAAAAACCGTGCTGGCGCTC
CAACACCGCCGCCTGCCGACCAGCCTGTTCTACCAGTCGCCCAATCCACACATCGACTTTGCG
GCGAGTCCGTTCCGCGTGAACGGCCAGACTTCGGATTGGGTCGCGCCAGAGGGGACGCGGTTG
CTGGCGGGAGTGAGTTCGTTCGGTATCGGGGGAACCAACGCCCACCTGATCGTCGAGGAGGCG
CCGAAAGCGCTACCGACGACAGCGGCACCTCTGTCGACGGAGCCGAATGACCTCGACGCGGGC
GACGCCGACGGGCTAGTGCTGCCGATCTCGGCCCGCACGCCGACCGCCCTGGCGCACATCGCG
ACCAACCTCGCCAATCACCTGGAACGACATCCGACCATCGCCCTGGCCGACGTCGCCCTGACC

CTTCAGCTGGGCCGTCGCCAATGGCCCCATCGCCACAGCCTGATCTGCCGGAATCGAACGGAG
GCGATCAAGCTGCTGCGCGCCGTCGTCCACTCCGCGGAGGTGCCGCCAGCTCAGGCGCCGGTC
TCGGATGCGCCGCGCTGTGTTTTTCTTTTTCCCGGCCAGGGCGCCCAATACCCGAGCATGGCC
CGCGACCTGGTTCGAAACTGTCCCGACTTCGCCCTGCACCTGGACCCCTGCCTCGACCAGTTG
GCCGAACTGCTTCCCGAAGATCCGCGTTGCATCCTGTTCGGCGATGGCCCCGCCGATCGGCTC
GACCAGACGGCCTACACTCAGCCGCTGCTCTTCTCCGTGTCCTACGCCTTGGCGCGCTGGTTG
GGCGATTTCGGCATTCGCCCCGATGCGATGATCGGCCACAGCCTGGGCGAATACGTGGCGGCC
TGCTTGGCCGGGCTTTTCTCGCTGAGCGATGCCCTGCTGCTGGTGAGTGAACGCGGCCGCCTG
ATGGGCTCGGCCGCGCGCGGAGCGATGCTGGCCGTCCCCTTGCCCGAATGGGAACTGGAGGAA
CGCCTGGAGCTTCTGGCCGACGACCGAATCAGCATCGCGGCGGTCAACACCGCCGAGAGCTGC
GTCATCGCGGGACCCAGCGAGGCGATCGAGCGCTGCGCCCAGCGCTGGGCCGCGCAAGGCCTG
ACCTGTACGCCGCTGCGCACGTCCCACGCCTTCCACTCCGCGATGATGGAGCCGATTGTCGAA
CCCTTCGGCCATGTCTTGGCACGGGTCACCTTCGCGCCGCCGCGCGCGCGCTGGATCTCGAAC
CTCGACGGCAAGCCGATCGATTCCGCGGCGGTGATGCAGCCCGACTATTGGGTGCGCCACCTG
CGCCAACCGGTCCGCTTTCACGAGGGACTCAGTCACCTGTTGGCCGAGGACACCCATGCTTGG
GTCGAAGTGGGTCCCGGCCGAACCCTGTCCTCCTTCGTCCGCCGCCACCCGGCCTACCGTCAC
CAGCCAATCGTCAACCCCATGCGCCATGCAGTCGAGTCGACGGGCGACGTGCGCCGGTGGCGC
CAAGCGCTGGGCGAACTATGGCGGGCCGGCATGCCGGTCGCCTGGGAGCGGCAGCGGCGCGGC
CGGCATGCCGGACGACGTGTGCCGCTGCCGGGCTACCCCTTCGAGCGGCGGCCCTTCGCGGCC
CGAAGACCGGTGGAGCTGGCGCAGCCCGCGCCCAAGGCGGAGCTGGTGAAAAACCCCGATCCC
GCGCGGTGGCTGTACCGCCGCGTCTGGCGCCCTGCCCAGGCTGCGGCCGGCGGACTGGCGGTG
CAGGCGACCGTTCTGGTCTTCGGCGACGGGTCCGAGCTGTGCCGCGCGGCGGTCGCTCAGGTG
CAGCGCCAGGGGCTGAAGTGCGTCTCGATCACCGCGGGCCGCCAATTCGCGCGGAGAGCGAC
ATGCGCTTCACGCTTGACCCCGCTGATCCGCGCCAGCTCGACCAGCTCTTCGCGGCCCTCGAT
GGCTCAGGCTCGCGGCCGCGGTACGTCCTGCACCTGCTGACCCTGAACCCGCCCCCGGATGCC
TCGGCGATCATCGCTCACAGCTACTACAGCCCGATGGCCTTGGCTCATGCCTTGGGCGCCCAC
GAGATCGCGCCTGTCTCGATCACCGTCGTCACCGCCGGGGTCGTCGCCGTCGCGGACGAAGCG
ATTCGCGAGCCGCTGCAGGCGCTGATCGTGGGCCCGTGCCTGGTCATCCCGCAGGAGTTTCCC
GGGCTCAGCGTTCGGCTGCTGGACGTCAACGTCGACGATCCGGCACCGCGTCTGGCGGAGCGG
CTCGTGGCCGAGCTCTCGGGCACGGATCACATGGTGGCGCTGCGCGGCGGCGAGCGCCTAGTG

GCCGATGTCGATCAAGTCGATGGCCTCGGTGTGGGGATCGCCAAGGTGCCCTTGCGCCGCGAG
GGCCACTACCTGATTCTCGGCGGCCTGGGCGATATCGGCTACCACTGTGCCCGCTATCTGGCC
CAAACCTACCGCGCCAAGCTGACGCTGACCGCGCGTTCGTCACTCCCGCCGCGCGCGTCGTGG
GAGCGAATGCTGCGCGAGGGAAACCTGGATTCCCGGCAGCGCACGCGCATCGAGCGCGTGTTG
TCGCTAGAGGCGTGCGGGGCCGAAGTCCAGACGGCTGCGGTCGACTTGGGCGATCGCCATCGC
TTGGCCGATGTGTTCCGCGAAGCACGGGGCCGATTCGGCGCCATCGCGGGCGTGATTCACTCG
GCGGGGATTCCGGGACACGTCCACTCGATCGACGAGCTGGTGCGCGTCCGCGACGAAGCCCAA
TTCACCGCGAAGGTTCGAGGGCTGCACCACCTGGCCGAGGTCGTCGATCCGCTGAACCTCGAC
TTTTGTCTGCTGTTCTCCTCGCTCTCGACCGTCCTCGGCGGGCTCGGCTACGGCGCCTATGCA
GCGGCCAACGCCTACATGGACAGCTTCGCCCGCCGCCACGATCGGCCGGACGAATGTCGTTGG
ATCGCGGTCAACTGGGACGCCTGGCTGTTCGAAGCCAAGACGTCGTCGGTCGGCGCCGAATTG
GCGCGCCTGGCGATCGTGCCCGAGGACGCTCCGGCCCTGTTCGCGCGGGTGCTAGAGCGACTT
CCGCAATCGTTCATCGTGTCCACCGCCGACCTTCGGGCCCGCATCGACACTTGGATCCGGGAC
AAGAACCGCGTCCCGCCCGCCGAGATCCGAGCGGTTCAACCGCGACCGGACCTGAGCCAGGCG
TACGCCCCGCCGATCGGCCCGCTGGAGATTCAACTCTGCGGGCTGGTCTCCGCCTATTGCCGG
TTCGACCGGATCGGGCGGGACGATTCCTTCTTCGAAATCGGCCTCAGCTCGTTCGACTTGATC
CAGCTCAGCTCGCGCATTCACCGCATCACCGGCAAGGATCTCAATACGACCCAACTGTTCAGC
TACCCCACCGTGCGCGCCTTGGCGCTCTTCCTCGGCGGCGAACCGGAGGGGCTCGCGGCGGAG
GAGCCCGCCATGGAGAACCTGTGGCTGCAACGAAGCGATGCGACCCTCGATGAG

(2) peptide sequence

Seq ID No 64 (>pEPOcos6_ORF12.pep)

MTVEHETGFEIAVIGLACRVPGAADVAAFWRNLVEAKESVRFFEDHELRAAGVPEEILRLPNY
VKAKPLLADGEAFDADFFGFHPREAAYLDPQVRLLHECCWTALEDAGYDPAQYAYPIGLFAGV
SSNLSFLFDRIDPRDSPLQKRYVAELNAASFATQIAYRLDLKGPAISIQTACSTSLVAIHLAA
QSLIGGECHMALAGGATLEVPKKPGYLYREGYINSPDGHCRAFDADAAGTIFGDGVGIVLLKR
YRDALRDGDHVYAVIKGSAINSDGHRKVSYTAPGKSGQVAVIRAALAAAQVEPQTIRFVEAHG
TGTLAGDPIEVEALTEVFAEAGRGTCALGSVKTNIGHLDVAAGVAGFIKAVLALERRVLPPSL
HFVRPNPAIDFNGPFYVCRQIERLTENGRLRAGVSSFGIGGTNAHVILEEAPAPEARLPAGSP
PGASPFLFPLSAKTPDALAGRCHDLADHLRAHPELLLADVALTLQMGRASFAYRHVVQAATAE

ELIRGLGAFRQESIRKRRNRVQWVLAGEAMSLDAGLRLYADWPVYRERVDVCLAIVAKLRQID
GRSFLHEWIERPREVPAEWSTALAFMFHCALAQALSQAGLHPQRMWSRGLGGQVGVVLAESLS
LEQALALVLCQTPVPGDATPQRERLVRTLEGCRFRPPRFLISADSSGRPLDLAEFAHVDFWCG
GQSASPNEAELRSWSDAAPELVTLAIGPSFLEAASGTVGLAIDPKRPMTCVQRTVAALWEWGC
DVRWAAFTSSTGRRVPLPTYPFVRVIPTIGDPLRGAGAEDDLIAASASASAGSPPEPSANSAA
ERPRAQSSIASATTPAPSHTSASVAVATILETVRAYFGFAAVRSTDAFFELGASSLDLVNLGQ
LLSDRLGREVPTLLLYDHPTPDQLALALTSAALSAEAPPLRGGHRASTSGTAASSAASTAPTF
PGDAHSQPSFVREQDIAIIGMAFRGPGADDLDAFWNNLVEGVESITFFSEDELLAAGVPREHL
ASTRYVRAKGELTGMMDFEPEFFGYSAREAAVMDPQFRVFHECSWHALEHGGYDPTRCAASIG
VYAGVTNHLPWLMRTLPHLTEEEQFGALLLTDREFFAPLLSYKVGLRGPAISLQTACSTSLVA
IGTACRELRAGACQMALAGGVTASIERCGYFHQEGYILSPDGHTRSFDAAAAGTVFGDGVGMV
LLKPLAQALADGDTIHAVIKGIGINNDGARKVGFTAPSRAGQTEAIRAALRDAGVASNRVSYV
EAHGTATRMGDPIEVEALTQAFRAEADGPLPPGSCLLGSVKSNVGHLNAAAGVAGLVKTVLAL
QHRRLPTSLFYQSPNPHIDFAASPFRVNGQTSDWVAPEGTRLLAGVSSFGIGGTNAHLIVEEA
PKALPTTAAPLSTEPNDLDAGDADGLVLPISARTPTALAHIATNLANHLERHPTIALADVALT
LQLGRRQWPHRHSLICRNRTEAIKLLRAVVHSAEVPPAQAPVSDAPRCVFLFPGQGAQYPSMA
RDLVRNCPDFALHLDPCLDQLAELLPEDPRCILFGDGPADRLDQTAYTQPLLFSVSYALARWL
GDFGIRPDAMIGHSLGEYVAACLAGLFSLSDALLLVSERGRLMGSAARGAMLAVPLPEWELEE
RLELLADDRISIAAVNTAESCVIAGPSEAIERCAQRWAAQGLTCTPLRTSHAFHSAMMEPIVE
PFGHVLARVTFAPPRARWISNLDGKPIDSAAVMQPDYWVRHLRQPVRFHEGLSHLLAEDTHAW
VEVGPGRTLSSFVRRHPAYRHQPIVNPMRHAVESTGDVRRWRQALGELWRAGMPVAWERQRRG
RHAGRRVPLPGYPFERRPFAARRPVELAQPAPKAELVKNPDPARWLYRRVWRPAQAAAGGLAV
QATVLVFGDGSELCRAAVAQVQRQGLKCVSITAGRQFARESDMRFTLDPADPRQLDQLFAALD
GSGSRPRYVLHLLTLNPPPDASAIIAHSYYSPMALAHALGAHEIAPVSITVVTAGVVAVADEA
IREPLQALIVGPCLVIPQEFPGLSVRLLDVNVDDPAPRLAERLVAELSGTDHMVALRGGERLV
ADVDQVDGLGVGIAKVPLRREGHYLILGGLGDIGYHCARYLAQTYRAKLTLTARSSLPPRASW
ERMLREGNLDSRQRTRIERVLSLEACGAEVQTAAVDLGDRHRLADVFREARGRFGAIAGVIHS
AGIPGHVHSIDELVRVRDEAQFTAKVRGLHHLAEVVDPLNLDFCLLFSSLSTVLGGLGYGAYA
AANAYMDSFARRHDRPDECRWIAVNWDAWLFEAKTSSVGAELARLAIVPEDAPALFARVLERL
PQSFIVSTADLRARIDTWIRDKNRVPPAEIRAVQPRPDLSQAYAPPIGPLEIQLCGLVSAYCR

FDRIGRDDSFFEIGLSSFDLIQLSSRIHRITGKDLNTTQLFSYPTVRALALFLGGEPEGLAAE
EPAMENLWLQRSDATLDE

pEPOcos6_ORF13 sequences:

(1) nucleotide sequence

Seq ID No 65 (>pEPOcos6_ORF13.seq)

ATGAAATACGAAACCACCGGATTGGAATTGGCCGTCATCGGTCTCGCTTGCCGCTTTCCAGGC
TCACCCGATCCCGAACAGTTCTGGTCGAATCTGCGCGCAGGTCGCTCCGGAATCCGCCATTTC
AGCGATGCCGAGCTGAGCCACATCCCCGCATCCCTGCGTCACCATCCGCATTACGTCAAGGCC
AAAGGCGCGCTGGACCACGCCGATTTCGAACCAGCCTTCTTCGGCTACTCGCCCAAAGAGGCC
GAGGTGATGGACCCTCAATTCCGGCTGCTCCATGAGTGCTGCTGGGAGGCGCTGGAGTCAGGC
GGCTATGCGCCGAGCCAATTCGCGGGTCGGATCGGCTTGTTCGCGGCGGCGGCCTTCAACGAC
GGATGGATCGCCGGTACCCTCGACCGGCTGCGCACCGGCGTGGGTTTGAGCTCCCTGGAAACC
GCGTTCTTGACCCTGCGCGATTACCTGACCACCCAGATCTCCTATCGGCTCGATCTGCGGGGC
CCCAGCCTGCTTGTCCAAACCGCCTGCTCGTCGTCGCTGGTGGCGGTCCAGCTCGCCCAGCAG
GCGCTGATCTCCGGCGAATGCGCCCTGGCCTTGGCTGGCGGCGTGTGCGCGACCGATCCGCTG
CATTCGGGATACCTCTATGAACCCGGCAACATCTACGCGCGCGACGGCGTCTGCCGACCGTTC
GACGAGGCAGGCGCCGGTACGGTCTTCGGCGACGGGTGCGGCATGGTCCTGCTCAAGCGGCTG
AGCGACGCCCAGCGCGACGGCGATACGATCTGGGCGGTCATTCGCGGGGCGGGCGTGAACAAC
GACGGGCACCACAAGGTTGGCTACACGGCTCCTGGCACGAGGGGCCAGGTGGCTTTGCTTAAA
AGTGTTTATCGCGCGAGCCGGGTCGACCCGGCGACGCTCGGCTACCTGGAGGCCCATGGCACC
GGCACCGCGCTCGGCGATCCAATCGAGGTCGAGGCGCTTACCCAGGCCTTCGCCAGCAAACGT
CGCGGCACCTGCGGCTTGGGCTCGGTCAAGGGCAACCTGGGTCACCTCAACACGGCGGCCGGC
ATCGCTGGACTGATCAAGGTGGTGCTGGCGCTGAAACATCGCGAAGTGCCACCCACCCTCAAT
CTGCGCCGTCCCAATCCGAAAATCCGCTTCGACGAGACGCCGTTTTTCCCAGTCGTCGAGTTG
CAACCCTGGCCAAGCGGACCGGCCCCTTGCGAGCCGGCGTGAGCTCCTTCGGCATCGGCGGT
ACGAACGCCCACGTCATCCTCGAGGAGGCACCGCCGACGGCCAACCCGGCGCCACACGGCAGA
TTCCGACTGTTGCCGCTTTCGGCCAAGACACCGGCTGCGCTCGAAGCGAAGCGCCGCGATCTG
GCCGGCTTCCTCGAACGCCACCCGGAGACCTCCTTGGCCGACCTCGCCTTTACCCTGCAACGC

```
GGCCGCGAGGTCTTCAGTCACCGCGCCTGCCTCGCCGTGGAGACCTTAACGTCCGCGCGCACG
CGGCTGAGCGGCGAGTCGTCGAGCACTTGCGTGGTGGGCCCCGCGCCCAGCGCCATATTTCTG
TTCCCTGGTCAAGGCAGCCAGCTCGCCGGGATGGGCCGCGGTCTGTATCACCATTTCGAGCCG
TTCCGCACGGCCGTCGATGCCTGTCTGCGCGAGCTGGAGCCAGGACTGCGGCAAGCGCTCAGC
GCCCATTTCGATCCGAATCGCGGCGCGGACCCACCCGATTCGACGACCTTCGTCCAACCCTTG
TTGTTCCTCGTCGAGTACGGGGTGACCGAGTGGCTACGCTGCTTGGGTGTGCGGCCAACAATG
GTGTTGGGTCACAGCTCTGGCGAGTATGCCGCAGCCTGCGTCGCGGGCGTTCTGTCGCCGTCC
GCGGCGGTCTCGCTGCTGGCCGAGCGCGAGCGGCTGCTGCGCGACCTGCCAGCCGGCGCCATG
CTCGGCGTCCCGCTGGCCGCCGAGGCGCTCGAGGCGATGTTGCCCGACGCTCTCGATCTGGCG
GCGATCAACGGCTGTCAGCTTTGCGCCGTGTCCGGGCCGGTCGCGGCGGTCCACGCCTTCAAG
GCCCAACTGGAAGCCGCCGGACATCACGCCCGCCTGTTGCACACCGATCGCGCCTTCCACTCG
CGGCTGGTAGCACCGGTGCTTGACCGGTTCCAGGCAGCCGTTCAACACGTGGAGCTGCGGCGG
CCGCAAGTACCTTACCTCTCGACCGTCAGCGGGCGATTGGAGGCGGATGGGCCGGCGAACCCG
CACTACTGGGTGCGTCACCTGCGCGACACGGTGCGGTTTGGTCCAGCCCTGGAGGCGCTGCCG
CCGGTGGATTCCTTCGTGTGCATCGAGGTGGACCAGGCTCGGCCTTGAGCACCATGGCGCGC
GAAACGTTGGGTTCCCAGGCGCGACTGATTTCGTTGCTGCCGCGGCCGCGAACGGGGCAAATC
GAGCCCGGTCCGGTATTCGAACGACTGGCGGCGCTTTGGCGCAGCGGGTTGACATTGGATTGG
TCTAAATTGACGGGCGGCGAAGAGGGTCATCGAATTCCCTTGCCAGTCTACCCGTTTCAGCGC
AGCCATCTGTCGAGCTCCCTGGCGGCGGGCCACACGCCTTCGTCGCGGCCTGCAGTCGAATCA
GGCGCCATCCTTGCCGAGCGATCCGCAGGGGAAAACGCTGAAACCCGGGATTGCCCGCTGCCA
ACCGCCACGCTCGAGCCCAAGGCGGTCGCTCCGGCCCCACTCGAGGCTACCGACGCCGCAGGT
ACTCGCGAGCGACTGGCCGAACTTTGGCGCGAGTTGCTAGGGTTGACCTCGATTGGGCCCGAC
GACCATTTCTTCGACCTGGGCGGCCACTCGCTGACCGCCACGCGGCTGCGCGCCCTGATTCAC
CAGCGGTTCGATGTCGATCTCGGGCTCGACGAAATCTTCGCTCATTCGCGTCTCTCCCAGCTG
GCCGCCCGTATCGAGGCGGCGGCCAAGAGCCGATTTTCCTCCATTCCCAGCGCGCCGGACCAG
GACGACTATCCCTTGTCATCCGCCCAGCAGCGGATTCACAGCATCGTCACGAGGGCCGAGGTC
GGCACTGCTTATAATTTTCCGATCGTCCTCGAGCTGCAGGGCGCTCTGGATCGAGTGCGATTC
GAGGCGACGTTCGCGGCATTGTTCCGGCGTCATGAGGGGTTCCGCACCCGCTTTGTGATGCGC
GATGGCGGGCCGCGCCAGCGCATTGTACCGGACGTGGCGTTTCGCCTGCCGCTCACCCAGGTC
GAGCCAGAGCAGGTTCCCGGGCGCATCGAGGCCTTCATCCGTCCCTTCGATTTGGAACGCGCG
```

CCGCTGTTCCGCGCGGAGCTGTTGCAGTTGGCCGAGCAGCGCCATCTGCTACTTTTCGACATG
CACAACTTAATTGCCGACGGTATCTCGCTCAACCTGTTCGTCGCCGATTTCGCGGCCCTGTAC
CATGGTCGTCCGCTGGCGCCGCTGAAACTCCGCTATCGCGACTATGCCGTTTGGCAAGAGGCG
CGGCTGGCCTCCGATGACCTGCGCAGCCAGCGCGAATGGTGGCACCGGCGGCTTTCGCCGCCG
GTCGCCACGCTGGCGCTCCCTCCCGATTTCCCGCGTCCGGCGGTGCGCCGCTACAAGGGCCGT
AATGTGGTGTTCCACCTGGACCGGGAGATCCGCGACCGCCTGGTGGCCCTGGCTCGAACCCAG
GGGGTCACCATGAACGTGATGATGCTGGCGCTCTGGGCTGCGCTGCTGCATCGCGAAACCGGC
CAATCGGAGCTGGTGGTCGGATCGCTGCTCGGCGGGCGGCCGCACAGCGAGCTGCATCCCGTG
ATCGGGCTCTTCACCAACTTTTTGCCCTTGCGGTTGGCGGTCGAGGGATCGACCCGCTTCGAT
CGCTTCCTTGCCGCTTGCCACCAGGTGTTTCTCGAAGCCTATCAGCGCCAGGACTATCCGTTC
CACTTGTTAGTCCAGGAACTCGTGCCGGTCAGGGACCCGTCGCGGTCGCCGCTGTTCCAGACC
TCGCTCGTCTACCACAACGAAATTGACGGCAAGACCAAGCTGGAATTGGAAGGGCTGAAAGTC
GAAGTGGTTCCCTTCGAAAAGGGTGTGGCGAGGCTGGATTTGAAGCTGGATGTGACACCTTTT
TCCGACCGACTCGAATGTGTTTTGCAATACGACTTGGATCTGTTCTGCGAGGAGACGATGCGC
GGCCTGATCGCGCGGTTCCAGGCGTTGGTGGCGGGGCTTGTCGCCGATCCGGCGCAATCGCTC
GCCGCCGCGAGCGTTTCCGGGAAGCGGGCGCTGCGCGCGGGCGTGGCCACGGCAAGCGAATCG
TCGCCGCAGTCACTGCCGCCGCAACCATCGACGGCGTACGCCACTCCCTCACCGCAGTCACCG
TCGCCGGTAGTCCTGACGGGACCCGCCGACCTGCCCGCGATCTTGGCGGCCTACGTGGGGCAG
AACCCCCATCCGTTCGCGATCCATCGGGGTCTCATTTTGGAGGCGCCGCTGGGGTTGCGAGCG
CTGCGGTCGGCGCTGGACGCAGTGCTCGGAGAACACACCCATTGGCGCAGCGTGCGTGCGGGC
GATCGCGCGCGGCGCGTGGATAAGTTGGAATTGACCAGCCTGGTGCGGCTCGACGACCTGCGC
GGGTTGGTCAATCCTCAGGCGAATGCCTTCACCCTGGCTTGGCGCGATCTGGCGATGCCGTTC
GGGGAGGGGCGTCCCCTGTGGCGACTCCGCCTGGCGTGGTCGGCTCCATCGCGCTGGTTGCTA
TTGCTGACGGTTCATCCATTGATCGGCGACAACGGCACGGTCGACCTCTTTCTGGCGGCACTC
GCCGATCACCTGCGCCGCGCGTCCGCTTTTCCCGTAGCACCGCTCGATGAGGCCGAGCTGGAG
GCGGAGCTGAAGTGGGGAGAGGAAGGGGAGGGCCTCGGGCTGACCGCGATCGCGCCGGTCCTG
GGCCAATTGCGCGAAAGTCGGCTGAGTCCTGTGGCCCAGATGTGGCTGGACGAGGTCTGTCGC
CGCCACGACCTCACCCCGCTAGAGGTCTTGGCGGCCCGGCTCCTCGATTGGACACGAAGCCAC
GGTCACGGGTCGATCGCTTTGTGGACGCCGCTGCCCGAGGACCATCCGCTTCGCGATGAAGGC
CGCTGCCTCCAGGTTCGCCTGCTGGAGGGGCCGCCGTCGCAGCGAGGAGCGGGCGATCCAAGC

TGGCTCGAGCAAATCGCCTTGAGACGGGGTACCCCTGCAACGGAGGTCGTTTGCCCTACTCCG
ACCCAACGGGCAGCCATCGACCTCGCGCTGGCCTGGCTGCCGCAGCCGCCTCTTCACGGTTTG
GTCGGAACCGTTCAGCCGTGGCCGGAATCTCCATTGGTCTGTCCGTTTCCCCTCAATCTCGCG
TTCCGGCCAAGCCATCCAATTGCCTACGCGCTCAAGCACGAGGCCACGCTCGCGGTCACGGCA
CGGGCGCGCGATCTGATGCGTTTCCTCGACGGCTTGGGCCCGGAAAGC

(2) peptide sequence

Seq ID No 66 (>pEPOcos6_ORF13.pep)

MKYETTGLELAVIGLACRFPGSPDPEQFWSNLRAGRSGIRHFSDAELSHIPASLRHHPHYVKA
KGALDHADFEPAFFGYSPKEAEVMDPQFRLLHECCWEALESGGYAPSQFAGRIGLFAAAAFND
GWIAGTLDRLRTGVGLSSLETAFLTLRDYLTTQISYRLDLRGPSLLVQTACSSSLVAVQLAQQ
ALISGECALALAGGVCATDPLHSGYLYEPGNIYARDGVCRPFDEAGAGTVFGDGCGMVLLKRL
SDAQRDGDTIWAVIRGAGVNNDGHHKVGYTAPGTRGQVALLKSVYRASRVDPATLGYLEAHGT
GTALGDPIEVEALTQAFASKRRGTCGLGSVKGNLGHLNTAAGIAGLIKVVLALKHREVPPTLN
LRRPNPKIRFDETPFFPVVELQPWPSGTGPLRAGVSSFGIGGTNAHVILEEAPPTANPAPHGR
FRLLPLSAKTPAALEAKRRDLAGFLERHPETSLADLAFTLQRGREVFSHRACLAVETLTSART
RLSGESSSTCVVGPAPSAIFLFPGQGSQLAGMGRGLYHHFEPFRTAVDACLRELEPGLRQALS
AHFDPNRGADPPDSTTFVQPLLFLVEYGVTEWLRCLGVRPTMVLGHSSGEYAAACVAGVLSPS
AAVSLLAERERLLRDLPAGAMLGVPLAAEALEAMLPDALDLAAINGCQLCAVSGPVAAVHAFK
AQLEAAGHHARLLHTDRAFHSRLVAPVLDRFQAAVQHVELRRPQVPYLSTVSGRLEADGPANP
HYWVRHLRDTVRFGPALEALPPVDSFVCIEVGPGSALSTMARETLGSQARLISLLPRPRTGQI
EPGPVFERLAALWRSGLTLDWSKLTGGEEGHRIPLPVYPFQRSHLSSSLAAGHTPSSRPAVES
GAILAERSAGENAETRDCPLPTATLEPKAVAPAPLEATDAAGTRERLAELWRELLGLTSIGPD
DHFFDLGGHSLTATRLRALIHQRFDVDLGLDEIFAHSRLSQLAARIEAAAKSRFSSIPSAPDQ
DDYPLSSAQQRIHSIVTRAEVGTAYNFPIVLELQGALDRVRFEATFAALFRRHEGFRTRFVMR
DGGPRQRIVPDVAFRLPLTQVEPEQVPGRIEAFIRPFDLERAPLFRAELLQLAEQRHLLLFDM
HNLIADGISLNLFVADFAALYHGRPLAPLKLRYRDYAVWQEARLASDDLRSQREWWHRRLSPP
VATLALPPDFPRPAVRRYKGRNVVFHLDREIRDRLVALARTQGVTMNVMMLALWAALLHRETG
QSELVVGSLLGGRPHSELHPVIGLFTNFLPLRLAVEGSTRFDRFLAACHQVFLEAYQRQDYPF
HLLVQELVPVRDPSRSPLFQTSLVYHNEIDGKTKLELEGLKVEVVPFEKGVARLDLKLDVTPF

SDRLECVLQYDLDLFCEETMRGLIARFQALVAGLVADPAQSLAAASVSGKRALRAGVATASES
SPQSLPPQPSTAYATPSPQSPSPVVLTGPADLPAILAAYVGQNPHPFAIHRGLILEAPLGLRA
LRSALDAVLGEHTHWRSVRAGDRARRVDKLELTSLVRLDDLRGLVNPQANAFTLAWRDLAMPF
GEGRPLWRLRLAWSAPSRWLLLLTVHPLIGDNGTVDLFLAALADHLRRASAFPVAPLDEAELE
AELKWGEEGEGLGLTAIAPVLGQLRESRLSPVAQMWLDEVCRRHDLTPLEVLAARLLDWTRSH
GHGSIALWTPLPEDHPLRDEGRCLQVRLLEGPPSQRGAGDPSWLEQIALRRGTPATEVVCPTP
TQRAAIDLALAWLPQPPLHGLVGTVQPWPESPLVCPFPLNLAFRPSHPIAYALKHEATLAVTA
RARDLMRFLDGLGPES

pEPOcos6_ORF13.1 sequences:

(1) nucleotide sequence

Seq ID No 67 (>pEPOcos6_ORF13.1.seq)

ATGACGCAAGCCTCGGCCGCGTCGACGTCCCAGGTCGCGCCGGAGGTCACCCCCGGCCGAAAG
GACGACGATGACGATCAAATCCGAGATGTCGGCCGTTGCTCACTCTGCGGAGAGCGGCTTCCG
CGCTGGGCCACGCGTGGGCGGCGCGATGAAGCGGGGCCGGACGCCGGAGCAGGCCGGCGTGAA
GCTGCTCCGCGCCCCGGTGAAGCGGAAGTGGCTGCCCCCGGCGCCCGTCCTGCGCCTGAGCGA
GCGGCGTATCCCGGAGGTGTGGGCAGGCTACCGCGCGAGCGCGGGATGACCCGAGCCCCGCCC
GCCGGCGCGACCATGACGCCGCCCCACGGGGCGAGTCGTCCGGCGCGCCGGCGCGCGTCGGGG
CTTCCGCCGCCGGGCGGGCAGGTGCAGGATGGTCGGGCATGG

(2)peptide sequence

Seq ID No 68 (>pEPOcos6_ORF13.1.pep)

MTQASAASTSQVAPEVTPGRKDDDDDQIRDVGRCSLCGERLPRWATRGRRDEAGPDAGAGRRE
AAPRPGEAEVAAPGARPAPERAAYPGGVGRLPRERGMTRAPPAGATMTPPHGASRPARRRASG
LPPPGGQVQDGRAW

pEPOcos6_ORF14 sequences:

(1) nucleotide sequence

Seq ID No 69 (>pEPOcos6_ORF14.seq)

ATGGTGACGCGTCCGACGTCCGACGGCATCGAGGACGAGCTCGCGCCGTTCCCCCCGGTCCTG
CGCGGCTGGCTCATCGAGGGCGAGCTCGGCCGCGGCGGGATGGGGCGGGTGTTCCGGGCGCGG
CACCCGAAGACGCGGGCGCGGGCGGCGATCAAGGTGCTGCTCGGCGACTACGCCCGCCGGCCG
GACGTGGTGGCCCGCTTCCGGCAGGAGGCGATCGCCGTCAACATCATCAACCACCCGGGAATC
GTCCGCGTCTTCGACTCCGGCGAGCTCGAGGACGGCTCGCCCTACATCGTGATGGAGTACCTG
GACGGCCGGGGGCTGCGCGACTGGGTGCAGGCCGTGCCGCCCGCGGAGCGGCCGCGGCAGGTC
GTGCGGCTCGGCTACCAGATCGCCTCGGCCATGGCCGCGGCGCACGCGTCCAAGGTCGTCCAC
CGCGATCTGAAGCCGGAGAACATCATGGTGGTCGAGGACGAGCTCGCGCCCGGGGGCAGCCGC
GTCAAGATCCTCGATTTCGGCATCGCGAAGGTCCTCTGGGGAGGTCTGCCCGAGGTGCTGGAG
CTCGAGGGGAGAGGCTCCCTCGCGCCCGCGTCCGCGTCCACGATCCGCACCGAGCTCTCGACG
CGGCCGGCGCCGACGGTGGGCGCCACGACCGGCCCAGAGAGCCCGCTGGGCGCGAGCGCCACG
CCAGAGAGCGCCCTGGGCGCGAGCGCCACGCCAGAGAGCGCCCTGGGCGCGAGCGCCACGCCA
GAGAGCGAGGCCCACGAGGAAGACGCGCTCCGGAGCCTCCCCGTCGTGACCAGCGGCAGGCCC
GCGATCCACCCCGCGCCGGTCGAGATCCCGCCCGAGGCGGTCTCCTCCGCGGCGTCGCGCGGG
TCGCGCGCGTCGATCGAGCCAGGCGCGCCCGCGCCGCAGAGCGAGGGCGCGGACAGCCCACG
ATGCCGTTCACGCAAGAGGGCGTGTGGGGCCTCGGGACGAGGAGCTACATGGCGCCGGAGCAG
GAGCGCCACTCCGGGAGCGTGGACGTGAAGGCGGATGTCTACTCGCTCGGCGTCATCCTCTAT
GAGCTGCTCGAGGGGCGGACGCCCGACGCGCCGAGCGCCGCGTGGCCGCCCCCGATGAGCGCC
GCCACGCCGCCCGATCTCGTCGCCCTCGTCCACCGGGTTCTGGCGTTCGATCCCGATGCGCGG
CCGCGCATGGCGGAGGTGGCGAGCGCGCTTCACCGGCTCGGCCGGGCGAAGAAGGAGCTCGAC
GAGGCGCTCTCGAGGTGGGTCGTCGGCGGAGGGGCGCCGGGGCTCTTGCCGTGCGGCTATGCT
CTTCTCGAACTGGTCCTCCTGGGCCCTGGGAACTTATACGATTCTTTCCAGCCTGTAAGTGCA
TTTTTCTTTCAATATCGTCCTCTCTTCATATACGAGGTGAGTTCTCTGAGGTCCTCCTATAAG
TCTGGGGTGTCCTATTCGGCCTCTTACTTGTTACTTCGCCTTCTTAGGAGTTTTTCCTTAATT
TTGCCCTCTTACATTCCCGTATTCATTCTAACTGGGCCCTATCTCATTCGC

(2)peptide sequence

Seq ID No 70 (>pEPOcos6_ORF14.pep)

MVTRPTSDGIEDELAPFPPVLRGWLIEGELGRGGMGRVFRARHPKTRARAAIKVLLGDYARRP
DVVARFRQEAIAVNIINHPGIVRVFDSGELEDGSPYIVMEYLDGRGLRDWVQAVPPAERPRQV
VRLGYQIASAMAAAHASKVVHRDLKPENIMVVEDELAPGGSRVKILDFGIAKVLWGGLPEVLE
LEGRGSLAPASASTIRTELSTRPAPTVGATTGPESPLGASATPESALGASATPESALGASATP
ESEAHEEDALRSLPVVTSGRPAIHPAPVEIPPEAVSSAASRGSRASIEPGAPAPQSEGAGQPT
MPFTQEGVWGLGTRSYMAPEQERHSGSVDVKADVYSLGVILYELLEGRTPDAPSAAWPPPMSA
ATPPDLVALVHRVLAFDPDARPRMAEVASALHRLGRAKKELDEALSRWVVGGGAPGLLPCGYA
LLELVLLGPGNLYDSFQPVSAFFFQYRPLFIYEVSSLRSSYKSGVSYSASYLLLRLLRSFSLI
LPSYIPVFILTGPYLIR,

or DNA sequences complementary to said open reading frames,

(b) DNA-sequences which hybridise under stringent conditions to regions of DNA sequences according to (a) encoding proteins or to fragments of said DNA sequences,

(c) DNA-sequences which hybridise to the DNA-sequences according to (a) and (b) because of a degeneration of the genetic code,

(d) allele variations and mutants resulting by substitution, insertion or deletion of nucleotides or inversion of nucleotide segments of DNA-sequences according to (a) to (c), wherein the variations and mutants offer isofunctional expression products.

10. DNA sequence according to any of claims 1 to 5, wherein the DNA is selected from the group consisting of

(a) the following DNA Sequence:

Seq ID No 71 (>Contig43)

CGGGTATTTGTGATATGTGGGCNGTAGTCGTATGCTTCATTAAGTACATC
CGTCCGTNGTAGAGAGTGACTCTGTCGCAGCGATAATAGACACGCTTGTG
ATGCTATAGGGAACATAGAGTCNTAGTAGATGATACGACGAGATATTNGT
ATAGAGCGTATAGACCGACGTGTGAGCGTCATAAGTGTTGTGTGTCATGA
GTGTGCTCAGAGGACGTGCAGACATTATATGAGCAGATGATGAGAGAGAA
TCAATGCTGCAAGNTATTCGTCGAATCTACATTATATCGAATCGTGTATG
TGCGTTTGTCGCAGCGCGATNCGATGAGATACCGAAAGGGTATGTATCTA
TNTTCGTGACGCTCGATNAGAGCAAATCCGCTACCGTGGAGATATCGTGT
ATCGACTCCATCACGATCAGTATCATGATACGTCAAACGAGTACACTCAT
TATTGATAACACACGTANGTGTGCATGCACAGTTATCGAGTGTATTGTGT
GCATGAGAGGTATAGGATNTATAGGCGAGCATATATATCTATATATATAG
GTTAAGAGTAGAANACTATGAAGATGCAGGAAGTAGTATCTCGCGGACAA
ACGGNGTACCTAGCGGGGTTGAAGTATTATCGACAGTGTATAACGACTCA
ACAGGGNTACGAGGTACATTGTATTTACAGTGGTTGGAAGGATTGCGCGA
GGAAAGGTAGTGGTACCGTGTGAGCTACGATGCTCGGGATAATGGTGATT
AGATAGAACCTTAGCGTTGCTAGATGAGTGAGTGGTGGTATGAGTAGAGT
TTTTGTTCTAGCTTTGTGTCCAGCGAGGATTCGTTCAGTCTGAAGGGTAA
GAGTACGTCCATCGCACACCCGACCGTTTTGAGGAGTTCTCGGTGCGTGG
TCAGTGGGGTTTGGAGAAGACAGAGTTGATTCATAGGGTTATCAAACGAG
TTATGTGGATAGATGGTAGTGACCCCATTTGAGTGAGAGTGTTGGCGTTA
ACANCAGCAGGATNTAT

SEQ ID No 72 (>Contig44)

TAGGTCTTTGACACCATGGGAGCTGCTACCGATGTTGCCGAGCACGATCG
CGCCGGCGCCGACGAGCGACTGCAAGCCGGCCGCGGCCATTTACGCCTGA
CGAGCGAGGTGGGCGAAGTGCTGGTGCGCGCCGTGCGTGTCGAGCGCGCC
CAGGTCCGCCGTTGCGCCGTCGCCGAGCAGTAGCGCGCCGTCGAAGACGA

TCACCGCGATCGAGGTCAGCGTCGTGGGGGCGAGGCCGAGGAGCGCGAGG
ATGCCGAGCACGACGCCGGCCGCGCCGTAGACGAGCTTGGCGCCCATGCC
GCCGCCGAGCTCGGTGCGCGTGTCCCACTCGACGGGCGGCGCGGTGCGGC
TCAACGCGCCGAAGCGCGAGGCGATCGCGCCGCCCTGCGCGATCAGCGCG
GCGCCGAATACGATCGTGGCGATCTGAGGTGAGCTCTACTGGCATGATCC
CCGTCAGCCCGAGGATGGTGAGGACAATCGTCGCGGCGCCGCACAGTACC
TCGCACGAGCGAGCCTCCGAGCACGACCTTCGGCGTCGTCTCGTCCTTTG
GTCTGCGTCGCGCGCCCGAGTGCGGCGTTATGTGGCTCTCCGGCTGTGCA
AACCGTTCACGTTCTTCCGGTCCTGGAGTCAGCATCGGCATGATTCCCCC
GTCCTGCGGTGAGGCCTTGTCGCGCTCACGCGCGCTCCGACTTGCACGTG
CTGTGCCGGGTTCTCTCGCTCAGGAGGCGCCTCTCTTGGTGGTGCTTGCG
TCCTGGTCCGTTTGCCCGCCTGTGCGGTAGGTTTCTTGAACCAGGTGACC
TTCAGGGACCCCTTGATGCGCTCCATCGTGTCCTATGTCGATCCTTCTCT
GACTTGTATGGGTCTCGAACCAACTACGCTTGATCAGGCCTTCGAAGGGT
CCTTTGGGAGATCGACTCTGGATCCATACCGGGAGCCCCTGTTCTGCCGC
TCTCTTAAGTTTCCCCTTCTGTATCCGTGTCGACCGGAAACGCTTTATCT
CTAATGCGCTCTAATTGCGTCTCTGCCACACGTGCGCTTCACTCTGGATC
TACTTCTTCTCCCTAGTCTTCTACCTCCGTACCCTTATTTGTTGGTTCTA
TTTATTTCTTTTCGCTTCACCTCGCGTCATTGTCGCCTAGTGTTCCTCCC
TCATATCGCCTTTGGTCTCCCTCGAGCGTACAGTCCTCTCTCTTCAGATG
CTTTCCGGCTCCTCTTCTGCTGGCCCCTTATCCTTTCTAATACTTC

SEQ ID No 73 (>Contig48)

ATGCGCCCAGGAACACCCCGGTGCGGCTGCCGTCGAGGGACTGGGGTGCG
ATGCCGGCGTCCTCGAGCCCTTCCCAGGTGACCTCCAGCAGCAGGCGTTG
CTGAGGATCGAGCGACCGCGCCTCCCGAGGCGAGGTGCCAAAGAACGCGG
CGTCGAAGCCGTCCACCGCCTCGGTGAGCAGTCCGGCCCAGCGCGGCACC
TCCTCGCTGGGATGGACGCCGACCAGCGCCCAGCGCCGGTCGAGCGGCTG
GACCGCGTCTCGGCCTGAGTCGAGCAGCTCCCAGAATGCCTCCGGAGTGT
CCGCTCCGCCGGGGAAGCGGCAGCCAATGCCTACGATGGCGATCGGCTCG

GTCCGCTCTTGCTCCAAAGACGCGTTCTTTTTCGCAAGCTTGTCCATGAG
CAGAAGGGCATGCTCAAGCTTCCCGGCATTCGTGGTCGCCATACTCCCTC
GGTCCCTTACTCACCAACGATCTGCGCGAGCTGCGCCAGCTTTTCGGCGA
GCAACGCGTCCTTCTGCTCGTCCGTCATGCCCCGCAGAGCCTCGAGATCT
GCGGCATCGTTCTCGAAGCTCTTCTCCCGCTCGGTGGCCGGAGCGTGGGT
CGCGCCGGCATTCGGAAACAGAATGTCTAGCAAGCTCCCGCTCAGAGCTG
CTACGTTAGGGTAGGTCCATAGCAGGGTCGCCGGCACGGTGATGCCGAGC
GCGGCCTCGATGCGGTTGCGGAGCTCCAGGCCTATCAGCGAGTCCATGCC
GAGATTGCTGAACGGCACGTGCCGCTCGATCCTCTCCGGCGGAAGGCGCA
GCCCCCGCCCCAACAGCTCGCTCAAGTGCTTCTCCAGAATCAACTGACGA
TCTTCGGGCCTGGCGCTCTGCAGCGCCTCGCGCAGGTTCGACGCGTTCGA
CGCGCCTCGGTCGGCGCGGTCACGCTCCTTCAGCAGCTCCGCCCACAGCG
CCAATCGGGCCGCGTTGGGATAGAACTC

SEQ ID No 74 (>Contig49)

ACCACCGCTTCACTCAGTATGTACTTTGTTATACTCGTCTTAGTACAATG
ATATAATACTCATGTGTATTCTTAATCTCGGGGAGANAAAATTGGAATAC
TGGACACCGTTGCCGCATGCNGACTCTAGAGATCCCCCTGCGACGGTATC
CCACGGCACCGGTATGGCCGGCGCGCGCTCCGGGGGTCAACGCCCCGTGG
TTGCCTTCACGACAACGCCGGTCGGCGGGGCGCCGTTCGATGCCGCGGG
CCCGCGCGCGGCGGCGCGTTATCCTGTGGAGCATCTGGAGGGCGCTCACG
CACCTGTCAGTCTAGTTCTGGCCCGCCCGGAAGGAGTCCGGGAGGCCGAA
GTTGAACCCGATGTAGAGCGCGATGAACGACGGGAGCACGCGCGCGGGGA
TGTGCAGCGCGGCGCCGATCGGCGTCGCGAACAGGACGAGCTCGCCCGGC
ATGCCGGGCACGACATACCCGAGCAGAAACACGATCGGCACCACGAGCGT
GAGCTCGAGCAGCGATATTTCATGACCGACCGCGCGGGCGGCGGCGCCCG
CCATGACGAACACGCAGATCAACGTGCCGTTGACGTTGAGCCAACCGCCG
AGTCCCACCACGAAGAGCCTGAGCTCCTGCGGCACCGCCGGATAACATTT
GCGGACGAGGTGCAGGTTGAGCGGCGTCGCCAGCGCCTCGCTGCACGAGG
CCCACAGCAGCGGATAGACCTTGAGCCAGTAGTTGACGAAATAGTCGCGC

AGCGAGAACTCCGGGGCGGCGGCCTTCATCCGCAGCAGGCTCGCCGCATG
GAAGACGAGGCAGGCAGCGCCGACGACCCCGGACACGAGCAGATAGGAGA
GCATGAGGTCTCCGGCGCCGAGCGAGGGCGCCCCCGCCGAGGCGCGCGCG
AGGTCCCCCCCGTGCACCTGCGCGGCGAGCTGCGCGGGCAGCCCGCGGAG
ATAGGCGCCGAGCCCGAACATGAAGAGCGGGACCAGGCACTGCACGGCGC
CTCCCGCGCGCTCCAGCGCGTCCGCGGCGCGCTCCAGCGCGCGGGCGACC
CGCGGCGCGCGTACGGCCGCGAACGACGTCACGATGCCGGCGTAAAGGGC
GAGGAAGCACGGGCTCGAGATGACCAGGCCCGAGGCGCTATAGAGGGTGC
GCGCGGCCTCGAACGGCGCGCCCGTGCTGTGGCTGGGGAGCAGCGGCAGC
CCGAACACGAGCCATGTGACGACGACCCCGAACAGGCACGCTGCCAGACG
CTTGAGGGCGAGCCAGCCCATGATGTACGCGAGCAGCCGCCCCGGGCGCC
CTTGCCGGTGCAGGCTCACGAAGGTCGGCACGAGGACGACGAAGATGACG
ACCGGCGCCAGCGTGGTGTACCAATGCAGGAGACCGTCCATGGCGCGGGT
CGACCACCGCGTGACGCTGGTCTCTCTGTCTGACTCGATCATGGCCCATT
CGCCTAAAACTAATGATCCGTTCTCAAATTGGTCAAAAAAAAGTTCCCTT
AAGACTGTTTTACTCCGGAATATTAATATATTTCTGAGTGTGAGGTGATG
TTAATCACACATTCTGATATTCTCAAGGGGAATCCGTGTCATTGTGAATA
CTTCTCTCTCTACAAGAGAGGTTATATATGGTCTCGAATATCTCGTCCGC
TCTTATATATATTCTCTTGTGATAATATATATCGAGTGTGGGTACTCAGC
TCTCTTGGTGTAATCTATAACTCGGCATCTCTCATAATACCTTATATATA
CACACTCTCTCGGTCATATCTCGCATAATAGATATATTTTATATGTTCCG
CGTTTTATCCGAGTGGGATACACTTTTTCTATATTTTCTTTGGTGTGACG
CGTGGCGTCGAGCCTTATTATTGATTTGGTAGTCACGATATTCTCTAGAT
GACATCATACAGATGCTCATAACTCGATAAACACAGGTCGTACACGACGA
GACTCTCACTCTCACTCTT

SEQ ID No 75 (>Contig50

TCCCCAGTTTCTCCTCTCTACGCNCACATCTCAGCAGGAAAAAANATAAT
GGAGAATCGTTGCGCTCTAGCAGCATCTATAGGATCCCCGCTGCTCTTCT
TCATGCACCTCGTGGAGCAGAAGTTCATCAACGCCTTCGCGATCATCGTG

GCGGTGAGCTTCCTGGCGTTGCTCCTGTCGCTCGTCGTCGCCGACGTCGC
GACGCGGAACACGTTCCCGCCCGCGCCTTTGCCGGCGCTGAGCCCGCCGG
CGCCGGCGCTGATCCCGCCGGTGCCGGGCGGATCCGTTGGGCCGTCGCCG
GAGCCGCTGTCGGTCGGCCGGTGATCGGTTGTGCGGGCGCCGTGCCTCGG
GCTTACTACCCCCTCTCGCGGGTGGGGATATGGCCGTGGATGAGGGAGGC
GATGAAAATCGTGATCGCCACGTGCGCGTTGTTCTAGATCGTCCCAGGCT
GACCGTCGGGAGCGCCCAGCACGAGATGAAGAGCCACACCGCGAGGACCG
TGTTGAGGTACCGCACCGCAGGGGCGAGCATGGCGGTGATCGCGAAGATC
ATGCAGAGCAGCCCGAGCACCCATGTGTTCGTCCGCTGCGCGTGGCTGTG
CGGCCAGATGACGGCCGAGATGAGGAGCCAGAACCCGAGGACGACGTTCA
CGATGCGCGCCATGAGATTGCCAGCTCGAACCATGCTCCCTCCCACCTCC
GATCATGGGACCGATCGGGTCGCCACGGATCGATAACGGGCGTCAGGAGA
CCGTCAATCGGCGAGCTCGTGAGCCATGGCGACAGCCCGCCGACCGCGCC
GGCGGGTCTCTGGCCTGCTGGTCGCCGTGCCGGCGGCGGCGATGGGCCTG
CCTCCGGTCGGGCCGCGCGGGCGCGCGGCGCTCCGCGCGAAGCTGGAGAC
GCGGGACATCGCTCCCTCGCCCCGCGCTCGGACGAGCGGCGCGAGCCACT
TCTCGACGGCCGAGCGGGCACTAAGCTTCCGTCATGAGGCTCGGCGCACG
GCTCACCACGCACACGTTCTCGGCCGGCGCCGCCGGCATCAGCTTCGTCG
TCCAGCCGATCCCGGGCTCGGACCAGCTGTTCGTCATTCCGATCCAGTAC
CTGCTCGCGGCGTCGCTCGCGAAGGAGCGAGGCGCGCCGCTCTCGAAGGC
GGCGTGGTCCCAGGTCCACCAGCTCATCTGGGGCGGCGGCGCGCTTCGCC
TCATGCTCGGCTTGACCCTAGGGCTGATCCCGCTGGCCGGCGCGTTCACG
AACGCGATGACGGCGTTCCTCACGACCGAATATCTCGGGTACTACGTGGA
TAGAGCCCTCGACAACCCGGACAATCCGCCTCCGGCCCTGTCGATCCAGG
ATGTCTTGGACGCCATCACCTCCTTCTTCACCGGGCGAGCGCGGTAGGCG
AGCGGTCCCTGGGTCGAGCCCACCCTGCGGCTCTAGGAGCCGAAGGGCGA
GCTCCTCGGGAGCGGCGCGGCGTCACCACCAGATTCGCCGGCGCTTGCGG
CCGGAGCGTATCGCGACCGCCGCCACCGCCGCCACGGCGAGCACGGTGAC
CGCGGCGGCCGCCGCGATGGCGACGCTCCGGGCCGTGTGCTCGGCCTCGC
CCTTGATGCCGCCCCACCTCGCCTTGACCGTGACCAGGTCGCGCGCGAGC

CGCTCCTTGCTGTGCTCGATCTGCTCGGCAGGCCCCGCGGGCGCCACGAA
ACCTGCGCCGGCGCTGACGTGACCTCGCTCGCTGAGGGGACTGGCGACCC
TCTCCGTGTCGAACCGGATTGCGCTGGATGGATCCATATGTCCCGCGCTG
CAATCGTTGCTCCCCGCCGGCACATCGGAGTGCTCGCCGGATCGCGCGGC
AGCGCCGACGCCGTACTTCCATAGGATAGCCCACCCCATCGGACAAGCCG
GCTCCTGACGGCGGGCACCGAATGTTCGCCAGACGGGCACAAGGCGCACG
CCGCGGACGGATCGGCCGCACTGGCACTCCAGAGCGCATCGACGGATGGC
CGACGGATGTGCAATGAGGCGCCCGCACGAAGCGAATTGTCCCGAATACA
GCGAAGAAATCTATAGCGATGCGAGCAGAAGGATATGTCTATGGGGGGCA
GTCAGAAACTGGGGACAGTCAACGCACATATTCTCTCCAANTGCTAACGA
CAGCGTGCGCAGAGGAAGTATCCTACTAGTGTAAGAGGGACATTCGATGC
GACCGCATAAACATTCAGTCTACAACGCGTGAGAGGATGGAACACCCCGC
CCCTCTGAAGGCTAGACAACCATGAATATGTGCAGAGGAAACACAGAATT
CCAAAGGTGAGAACATATGTAGGATCGCGCCACCCGAGATTGAGTGAAGA
TATACATATATACTTATATGGATCTACAACATGGCGAACCGAACGTAGCA
NAATAGTAGATATAATTGTAATACTGAGCTACCGACAGAAAGATACACAC
GAGTGTACACACATCACACGCAGAGTGGTACCAAATTCACACCATGCGAG
CCACAATGTGACACGGAGGAGCACAGCATGGGCGCCACTATGGAGGAGAA
ACTACTGCAACCCACATCTGATGGACTGACCGCACGGACGGGACGTGTCT
ATACATACAGATACATCNGATGGAGGAAGATGCATGTGCGATGATATCAT
CGTCGCAAACTCATATGTCGAAGAAGATATGNGTCAACTCAGCACTACTC
ACACGATACGTGAACAGGAGTGACTAGGACATCNCATGGTGTGTCGGCGC
GTGCACGTGATATCAAACTCTCTGATCAACCACACACTATATAAGGAGTA
TCGAGCGGCGATGGAACACCCCCTCACAGCATACGTATATGCACAACGTC
TGAACACTCTNGAGACACAGTGGAAGG

SEQ ID No 76 (>Contig51)

GATCCAGTTACGCCCCGCCGCCTCGGTCACGCCGGGGTTTTCGGCGTCGA
CCGGGGACGGTCGGGGCGACAACCGGGGGTTGTCGTCAGCGGTTCGCGTG
GATGTGCGCGACGAAGTTGTCCGCCGCCGAGTCGTCCTGCGTGCCGCCGC

GCTGCCGTACGAGCAGCCCGAGGCACCGCGCGAGCTCCTGCCGCTTCCGA
TGCGCGGCGGCCTGCTCGGCCTCGTGCTCCGCGATTCGGCGCTGGCGCCT
GCTCTCCTCGGCGCGTTCCTGCGTCCGCGCGAACAGCACCTCCTGCACCT
CCGCGAGTTCCTCGGCGGTAGCGACCCGCACGTGATCCGCACGTCCTCCG
AGACCGGTGACGACGTCAACCGCGGAGTAGTACGCCTCCGCGCGGACGAC
GTCGTCGAGGTCGAGTTCAGCGTCGGCTCCGACGACCAGGCGCATCGCGT
CGGAGGTGAGGACGACGCCACCGTCGAGCGGCGCGGCCTCCGGCTCGTAG
ACGTACTGGAGGTCGGTCTCCTCGTCGTCGTCACCGTCGAGGTCGACGTG
CCGCCGCAGGCCTCGGGTCCACTCGATCGCGCGTCGTCCGGCGAGGGCTT
CCTCGTACTGGGCCCACCACGCGCGCAGCTGCTTCGGCGTGCCGTAGCCC
TCGGCCATGTCGGGGTCGAGCCCGGCGACCTCGATGTCCCACAGTCGGTA
GAGGATCTGGAACGGCGTCATGGACTTCCGGCCCCGGCCGGTCTTGGAGT
CCAGACGGGCGGTCTCCATCGCAGCTGCGCCGGCGGCTTCGAGGTCCTGG
TCGACGGAGTCGGGCCGCTCTCGCTTCCCGTCCTGGTTCTTGGTGAGGTA
CTCGATCAGCGCGACGTCGTCAGCTGACCGGACGATCGAGACCATCACGC
CGTGGCCCTTGCCCTTGCACTTGCAGCCGGGGGTGTCGCAGTCGGTCGAG
GGCTCGAACTTGGGGTCAGCCCGCTTGAGGGCGCCCGCCCACATCTCCCG
GAGCCAGTCCTCCCAGTCCCCCAGGTCCGTCTCGGAGGGCTCAAAGTGTC
CGACGACGTCACCCTTGGCCGGGGTGCCAGAGAGCTCGCCGCCGAGGAAG
ACCAGCAGGTTGAGGTGGGGGTGGTAACCGTTCTTCTTGGACCGGGTGAC
CTCAGCCGCGCGGACCATGCCGATGTAGCCGATCCGGTGGCGGATGCCGT
CCTCAGCGGGACGGACGTACTGCGTTCCGTCCTTCCGGGTGCGGCGGGCC
TCAGGGCGGCCGTAGAAGGCCGGGGCCGTGAGCATCCGCTGGTAGGCACC
GGGCGCGCGGCGGGGCTTGCCCGACCGGTCGAGGACCGGGGCGCCCTTGT
CGTCCAGGAGAGGCCCGCCCCAGAGCGCGGCGACCAGGCTGTCGAGGTCG
GTGGTCTGGTTATGCCGGGCGGTGAGGACGACAACGGCGAGCGTGCCGCC
GGCGGCGAGGTGCCGCAGAGCACCGGTCTTGATCTCCTCGGTCCGGCCAC
GGCGGATCGCGGAGGAGCACTCCGGGCATAACCAGATCCGCCCGCAGCGG
ACCAGGCCGATCGTGACGACGTACCCACGGCTCGACTTCGCGTAGATCAC
GCCGGTGTCCGGGTCGAGGACCCGCCGCCCGCACCCGCCGCAGGCGTCGA

TCCCGGAGACCCGGTTGAGCACCTTGCGGCCCTGGTAGCGGCGTACGGCA
GCGGTCGCCGCGCGCCTTGTGGTCTGTTCCGAAAGGGCTGCCGCCCTCTC
GGACTCTCCCGTTCCTCCCACGACTGCCACTTCCGCAAAGTCGCTGGTCA
GTGGGGGGTGGGAAAACTCTGTCAACCCTTTACCTAGGCGTCCCTTTTTG
CCAGGGGCGGTCTCACGGGCGGCCTCGGCGGCTCGGTCGGCGGCCTTCCG
GGCCCGCGCGGCCCGCTTCTTGCACGCCTCGGAGCAGAACTCCTTGGCTC
GCTTGCCGGGGGTGATCGTGAGGGCGGCGCCGCAGCGGCAACGGGGGCCG
GCGGGGACGCGGGCGGGCGACTGAGTCGGCGCGCCGATCAAAGAGGGGGT
TGCGGACGCCAAAGCGTCCCTTACGCTGGACACAGACGAGTACCTTGGTT
GGTAGCCGGGTGGACGTCAGAAGCGGTCAGGGATTAGGACCCCTGGCCGT
TTCGCTTTTTCTGGAGTTGTTCGGGTAGATCCTGCCGCATCGCCCGCCTC
ACGCGTGGCTCGCCGCGCGGATGCCTCAGAGGCCCCACCGGTCGTCAGGA
CGCAGACGTCGGCGTGCTCCTGGTGGTGAGTCACCAGCTCGACCACACGG
GCGCGGCCGGCGACGTCTC

SEQ ID No 77 (>Contig52)
CGGGATCTGGCCTTCATTAACCAACGACGGGGCAAACATAATAGGCTGGG
CATTGCGCTTCAGCTCACCACAGCCCGTTTTCTGGGAACATTTCTGACGG
ATTTAACTCAGGTTCTGCCTGGTGTTCAACATTTTGTCGCGGTACAGCTT
AATATCCACCGTCCAGAAGTTCTCTCCCGCTATGCTGAACGGGACACTAC
CCTTAGAGAACATACTGCATTAATTAAGGAATATTACGGCTATCATGAAT
TTGGTGATTTTCCATGGTCTTTCCGCCTGAAGCGTCTGCTATATACCCGG
GCGTGGCTCAGTAATGAGCGACCGGGTCTGATGTTTGATTTTGCCACTGC
ATGGTTGCTTCAAAATAAGGTATTACTGCCCGGAGCAACCACACTAGTAC
GTCTCATCAGTGAAATTCGTGAAAGGGCAAATCAGCGGCTGTGGAAAAAG
CTGGCCGCACTGCCGAACAAATGGCAGGCAGCTCAAGTGATGGAGCTTCT
GGTCATTCCGGAAGGTCAGCGTGTATCAGCACTGGAACAGTXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXGCTGGAACGATAT
ATCCGATTACGAAGTCTTGAGTTTTCCCGACTGAACTTTTCCGGTCTGCC
TGCCATTCAACTGCGTAATCTGGCTCGTTATGCTGGCATGGCGTCGGTAA

AATATATCGCTCGAATGCCACAGCAGAGAAAGCTTGCTGTACTTACTGCA
TTCGTTAAAGCACAGGAAATAACGGCATTAGACGATGCCGTTGATGTGCT
TGATATGCTAATTCTGGACATTATCCGCGAAGCAAAGAAAACCGGGCAAA
AAAAAAGACTCAGGACACTGAAAGATCTTGATCAGGCCGCATTGTTACTG
GCGCGGGCATGTGCATTGTTGCTGGATGATAATACAGATGTCCCAGATCT
CAGGCAGGTTATCTTCAAGTGCGTACCCAAAAACAGACTGGCAGAATCTG
TAAGCAAGGTTAATGAACTTGCTCGTCCACAGAACAXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXAAACGTTTTCTTCCGGCGGTGTT
GCGGGACCTGCATTTCCGTGCGGCACCGGCAGGTGAACATGTACTGGCTG
CGATTCATTATCTGGCAGAACTGAATGGTTCGAAAAAGCGCATCCTTGAT
GATGCGCCTGAACATATTATCACCGGTCCCTGGAAACGCCTCGTATACGA
TGCGGAGGGACGGATACAGCGTGCAGGTTATTCACTATGTTTGCTGGAAC
GCCTTCAGGATGCACTGCGCCGCCGGGACATCTGGCTTGAAAACAGTGAT
CGCTGGGGAGATCCTCGCGAGAAGTTGTTGCAAGGTGAAGAGTGGCAGAC
TCAGCGTATTCCTGTCTGTCGGGCACTGGGACATCCTGTCGATGGACGTA
AAGGTGTGCAACAACTGGCTATTCAGCTGGATGAGACCTGGAAAGCCGTG
GCATCACGATTTGAAAAGAATGCGGAAGTTCATATCTGTAATGAAGGTAA
ATATCCATCCCTGACTATCAGTTGTCTGGAGAAACAGGAAGAGCCACCAT
CATTGCTTCGTCTAAATAATCGGATCAAACAGCTACTCCCACCGGTAGAT
TTAACGGAACTGTTACTTGAGATAGATGCCCAGACAGGATTTACACATGA
GTTTGCGCATGTCAGAGAATCTGGTGCTCGAGCGCAAGATTTGCACATCA
GTTTATGTGCGGTATGAATGGCTAAGCCCTGTAATATGGGCCTGAACCCG
TTGATAAAGCACAATATACCAGCATTGACCCGCCATCGGCTCAGTTGGGT
GAAACAGAATTACCTTCGTGCAGAAACGCTGGT

SEQ ID No 78 (>Contig53)
ATTCCACGCGCTCACGGTCAGCTTCGACCCGCGCGAGCGCCCGGCGGCCG
CCTCGCAGAAGCGCGCGGTCACGCTGTCCGAGCTCGGCGCGGACGCGCAG
GCGCCGGAGTGGCCGTTCCTCGTCGGCGACGAGGCGGCGACCCGCGCGCT
CGCCGAGGACCTCGGGTTCCGCTACGCCTACGATCCGACCACCGATCAGT

ACGCCCACCCGGCGGCCGTCTTCGTCCTGACGCCGGACGGGCGGATCTCC
CGGTACCTGTACGGGACGGAGTTCCCGGCGCGCGATCTCCGGCTCGCGCT
CCTGGAGGCGAGCCGCGGCGGTATCGGCACGATCGTCGATCGGGTGATCA
TGACCTGCTATCGCTTCGACCCGGCGAGCCGGAGATACGCTCCGTTCCTA
CTCGGCTTCCTCCGGCTCGGGGCGGCGGCCATCCTGATCACGGTCGGCGG
GCTGCTCGCCGTCCTGTGGCGGCGCGAGCGCCGGCGGCCAGGTGCTCGCA
CGAGCGCCGCCGTCGGTCGTGACGCCGTGGCCGACCGCCAGGGGAGGTCA
CCATGATCAACGAGCTCCTGCGCAAGCTTCTTTTTCTGTCCGGCCAGTGG
TCGACGATCGTGTTCGACATTTACAAGCTGCTTTACTTCGTGATCTCGGT
GACGATGGCCGGCGCGACGCTCGTCGCCCTGTTCGCGGCCTACCTGATGA
TCCGGTACCGCAGGCGCCAGCGGGATGTTGAAGGCCCGTTCCCCGGAGCG
ACCGCGAGGCCTCCGCTCCTCCTCGAGGTCGGCATGGTGCTGGGCCTCAT
CGTCCTGTTCCTCGTCTGGTGGGTCATTGGAATGCGGCAGTATGCAGAGC
TCCGCGTCGCCCCCGCGGACCCGGTCGTGGTGTACGTGACCGGGAAGCAG
TGGATGTGGAAGTTCGCCTACCCGGAGGGCCCGAGCTCGGTGGCGACGCT
CTATGTGCCGGCGCGTCGGCCGGTGAAGCTCGTCATGACGTCCCGGGACG
TGATCCACAGCTTCTTCGTCCCCGATTTTCGCATCAAGTACGATGTCGTC
CCCGGCCGCTACACCACGCTGTGGTTCGAGGCGACCGCGCCGGGCGCCTA
TCAGATCCTGTGCACCGAGTACTGCGGGACGAACCACTCCACCATGCGCG
GCGAGGTGATCGCGCTCGAGCCCTCCGATTTCGCGCGGTGGCTCTCCGAC
CGCGGGCGGGGCGCCGGTATCGCCGGACAGGAGTACACGCCGCCGTCGAC
GCCGGGCGAGGGGATCCCGCGCGAGCCGCTCAGCCTCGTCCGGCTGGGCG
AGAACATCGCGGCCGAGGAGGGCTGCCTGCGCTGCCACACGCCGGACGGG
ACACCGCACATCGGGCCGACCTGGGCCGGCCTCTACATGTCGGTCGTCCC
GCTGGAGAGCGGCGGCGCCGCGGTCGCCGACGACGCGTACATCACCGAGT
CGATGATGGATCCGCTCGCCCGGATCCACCGCGGCTACCAGCGGGTCATG
CCCTCGTTCCTCGGCCGGCTCCAGCCGGCGCAGGTCGCCGCCATCGTCGA
GTACATCCGGTCGTTGAGGGGCGTCGCGCCGGAGCCGGGCGCGCGGACGC
CGCTGCCCGAGGGCCCGCCCTTCCTGCGCTCCGGCCCGGAGCGCCCCGCC
CCGCTCAGCGGGGGCGCGCCGGTCGGCCCGATCGAGGGCGGCAAGCCCGG

GGAGGAGCTCCGATGAGCACGGAAGCGTACGAATCTCTGCCCGACGCGCC
GGCCGAGAGGCCCGAGCCCCGACTACCTCCATGTTTACCGCGGGGTGACG
GAGTGGCTCACGACCACGGATCACAAGCGGATAGGTCTCATGTTCTACGC
CGTCATCGTCGGGAAAGCTTCTTCCTCGGAGGCATATTCGCCCTCATCAT
GCGGACCGAGCTCCTCACGCCCGAGCGGACCATCATCGACGCGGCGACCT
ACAACCGGATGTTCACGCTGCACGGCGTGATCATGGTCTGGCTGTTCATG
ATCCCGTCGATCCCCAACGCGTTCGGCAACTTCGTCCTGCCGATCATGCT
CGGCGCCAAGGACCTCGCGTTCCCCCGGATCAACCTCGCGAGCTTCTACA
TCTACCTCCTCGGGGCGGCGATCGCGATGGCGGCATGATCGCGGGCGGC
ACGGACACCGGCTGGACGTTCTACCCGACGTACAGCCTGAAGACGCCGAT
GACGCTGTTCCCGGTCGTCTTCGGCATCTTCATCGTCGGCGTCTCGTCCA
TCATGACGGCGGTCAACTTCATCGTGACCACGCACACGATGCGCGCCGAG
GGGCTCACGTGGAGCCGCCTGCCGCTCTTCGTCTGGAGCACCTACGCGAC
GAGCATCATCCTGCTCTTCGCGACCCCGGTCCTCGGGCTCTCGATCCTGC
TCATCGGCATCGACCACGTGACCGCGCTCGGGATGTTCGATCCCCGGTTC
GGCGGCGATCCGGTCCTCTTCCAGCACCTCTTCTGGTTCTACTCCCACCC
CGCCGTCTACATCATGATCCTGCCGGCGTTCGGCGTGGTGAGCGAGGTCG
TCTGCACGTTCGCGCACAAGCGCCCCGCGTCCTACTGGGCGATCGCCATC
TCGTCGCTCGGGATCGCGTTCGTGGGGTTCTGGACGTGGGGCCACCACAT
GTTCGTGGCGGGGATGAGCGAGTACGCCGCGGACGTCTTCGGCGTGCTCT
CGATGTTCGTGGCCATCTTCTCGGCCATCAAGGTCTACACGTGGGTCGCG
ACGCTGTACAGGGGCTCGATCCACTTCAACACGCCGCTGCTCTACTTCAT
CGCCTTCCTCTTCCTGTTCGTCTTCGGGGGGATGACGGGCGTGGCCGTCG
CCACGCAGTCGCTGGACGTGCACTGGCACGACACATACTTCGTTGTGGCG
CACTTCCACTTCATCATGGTGGGCGGGACGCTCACCATGTTCCTCGCGGC
GGCGCACTACTGGTTTCCGAAGATGTTCGGGCGCCTCTACTCGGAGCGCG
TCGGGCTCCTCTCGGCCGCGTCGGTGTTCCTCGGCTTCTTCTTGACCTTC
TTCCCGCAGTTCCTCCTCGGGAACATGGGGATGCCCCGCCGCTATTACAG
CTACCCGCCGCGCTACCAGTGGCTCCACGTGCTCTCGACCGGCGGCGCCT
ACCTGCTCGCCGCGGCGCTCGTGATCTCGCTCCTGAACCTCGTCATCGCG

```
CTCAAGTGGGGCCGGAAGGCCGGGAGGAACCCCTGGGGCGGGCGCACGCT
CGAGTGGATGACCGGCGAGCCCTTGCCGCCCAAGCACAACTTCCCGGTCG
CGCCGCTCGTCCGCCGCGGCCCGTACGAGTTCCAGCTCTCCGAGGAGGAC
GCCCGTGCGACAACCACGCCCGCTGCGTGAGCAGTTCGAAGATCTCGAGA
AGCAGACGCACGCGGCCCGCCTCGGGATGTGGTTGTTCCTCGGGAGCGAG
GTGCTCCTCTTCACCGGGCTCTTCGCGCTGTACGCGGCGTACCGCGAGCT
CTACCCGCGCGATTTCGCCGAGGCCATCGCGCACAACAACGTCGCGATCG
GCACCACCATGACGCTCATCCTGATCGGCAGCAGCTTCACCGTCGCCATG
GCGGTGCACGCCGTCCGCGCCTCCCACCCGCGGCGCGCCGCGCTGTTCCT
CGCGGTGAGCGTGGCGATCGGGATCGTGTTCCTCGTGCTGAAAGGGATCG
AGTACGCGCAGCACTTCCGCGAGGGCATCTTCCCGGCCGGCGCCTACCGC
TTCGCGGAGCTCCCGACGTTCGGCGCGCAGATGGCGTTCACGCTGTACTT
CGCCATGACGGCGCTCCACGCCCTGCACGTCGTGGGCGGGGCCGGCCTCC
TCACGGGGGTCGCGTGGGGGTGCTGGAAGGGCCGGTACTGGGCATACGAC
CAGACGCCGGTGGAGCTCAGCGGCCTCTACTGGCACCTCGTGGACATCAT
GTGGATCTTCATCTGGCCGCTCCTCTACCTGACGCGCAAGTGACGGCACA
GCCCGGAGACGACCATGCCGCAAGAGCACGTCGCGGAAAGCACGCCCTGG
ACCCGTTACCTCATGGACGCTGATGGCCCTCATTGCCCTCACGCTCCTGT
CGTTAGCGCTCTCGTTCTTGCGCACGGGGGCTTGGGAAATACCGATCGCG
CTGCTCCTCCCCGTGGTGAAGAGCGTGCTCGGGCTTGCTTTCTTCTTGCT
CCGTCGAGGTGCATTAAGGTCATCAAACGCCTTTTTCAAACATTGGCTGC
TGGTGTACTTACTTGGACGTTTCATACTGTCTCTTTTATGGCCCCCAACG
GTCATTACGTCGACCTACACAATTTCTCTTCCCGCACCCTCCTACCTGTA
TCTCTAAGCACTGCCTTGCGTCCTGCTCTATTACATTCTACTCCGGCTGT
CCATGTGTGGGATTATATGCGCGAGGTACCTATTCCGCCGTGGAGTCTCC
ATTTACCTCTTGGACCTTGCCCGTTCTGAATCTCGATCTCCTCATGCGTT
GGTCCACCATGTATTACCTCCTAGAATCTTATACTCCATATCTCTATATA
TCTAGTTGTGCGTGTAATTGTGTCATATATTATCGCCACTGCTGTATGAA
TACCGTGCCGACGTGCTATATACGAAANTACTCCTCGGTCGATATCTCCA
CCTCATATATACCTCCGAGTGTAGTATACGCACGAGTGTATATACTCTTC
```

CTCTGGTCACGCGACTTCGTGCTGATATGATACCATCGTTCCATGTTACG
CGAAGTTACTCATAAGATCTCCTCACACATCAACGAGTGTACTCCTATGT
GTTTCATACAAACTCGATACCCTTCAGAGTAGTGTCATGCCTATGTGGTA
TGCATAATGTTAGTATACTTT

SEQ ID No 79 (>Contig54)
TGGGAAAGAGGGCCACAGGGGATGTAGCAGGACGCTTAATAGTAAATGAC
GAGGGTGTGCCGACGAGACCCGTAGGAAACAACGGGCACAGACGAGAGCA
ATAAAGGGGGTTGGAAGGTACCCCGGATAGAGTAGAGAAGGCTAGCGGAC
GAGTAAGACGCGGAGGAAATAAGTCGGCGTCGTAGAAGTTCTGTGGAGAA
GGTACGACTCTTAAAGACCTAGGCGGGAGACAGTTTCCACCCGAGGCAGA
GCAAGACCACAAGATTCAGAGGGAGTAAGGAGTTCCGAATTGGAGAGGTT
GAGGGGCGTGTGAGCCGTCAAGTGGGGCGCGTACGCAAAGAAAGAAGCGT
CCATGTCAGAGGCCCAGCGGCCGTTGCGGCCCTACATGGGTGATCACGGT
GGGCCGTAGGCGGACCGGAGATGAGCGCGGCTCCGCCCACCGGACGGCGA
GGGGCACGGCGCCTTCGTCCGGGGCGCTCGCGCGCGATCTCGCGCGCGCG
GGGGGCGCCGGCGTCGCCGTCGTCGGCGACGGCCAGCCGCCCATCGTCCA
CGCCCTCGGGCACGTCATCAACGCCGCGCTCCGCAGCCGGGCGGCCTGGA
TGGTCGATCCTGTGCTGATCGACGCGGGCCCCTCCACGCAGAGCTTCTCC
GAGCTCGTCGGCGAGCTCGGGCGCGGCGCGGTCGACACCTTGATCCTCCT
CGACGTGAACCCCGTGTACGCCGCGCCGGCCGACGTCGATTTCGCGGGCC
TCCTCGCGCGCGTGCCCACGAGCTTGAAGGCCGGGCTCTACGACGACGAG
ACCGCCCGCGCTTGCACGTGGTTCGTGCCGACCCGGCATTACCTCGAGTC
GTGGGGGGACGCGCGGGCGTACGACGGGACGGTCTCGTTCGTGCAACCCC
TCGTCCGGCCGCTGTTCGACGGCCGGGCGGTGCCCGAGCTGCTCGCCGTC
TTCGCGGGGGACGAGCGCCCGGATCCCCGGCTGCTGCTGCGCGAGCACTG
GCGCGGCGAGCGCGGAGGGGCGGATTTCGAGGCCTTCTGGGGCGAGGCAT
TGAAGCGCGGCTTCCTCCCTGACAGCGCCCGGCCGAGGCAGACACCGGAG
CTCGCGCCGGCCGATCTCGCTCAGGAGCTCGCGCGGCTCGCCGCCGCGCC
GCGGCCGGCCGGCGGCGCGCTCGACGTGGCGTTCCTCAGGTCGCCGTCGC

TCCACGACGGCAGGTTCGCCAACAACCCCTGGCTGCAAGAGCTCCCGCGG
CCGATCACCAGGCTCACCTGGGGCAACGCCGCCATGATGAGCGCGGCGAC
CGCGGCGCGGCTCGGCGTCGAGCGCGGCGATGTCGTCGAGCTCGCGCTGC
GCGGCCGCACGATCGAGATCCCGGCCGTCGTCGTCCGCGGGCACGCCGAC
GACGTGATCAGCGTCGACCTCGGCTATGGGCGCGACGCCGGCGAGGAGGT
CGCGCGCGGGGTGGGCGTGTCGGCGTATCGGATCCGCCCGTCCGACGCGC
GGTGGTTCGCGGGGGGCCTCTCCGTGAGGAAGACCGGCGCCACGGCCGCG
CTCGCGCAGGCCCAGCTCGAGCTCTCCCAGCACGACCGTCCCATCGCGCT
CAGGAGGACGCTGCCGCAGTACCGTGAACAGCCCGGTTTCGCGGAGGAGC
ACAAGGGGCCGGTCCGCTCGATCCTGCCGGAGGTCCAGCACACCGGCGCG
CAATGGGCGATGTCCATCGACATGTCGATCTGCACCGGGTGCTCCTCGTG
CGTCGTGGCCTGTCAGGCCGAGAACAACGTCCTCGTCGTCGGCAAGGAGG
AGGTGATGCACGGCCGCGAGATGCAGTGGTTGCGGATCGATCAGTACTTC
GAGGGGGGAGGCGACGAGGTGAGCGTCGTCAACCAGCCGATGCTCTGCCA
GCACTGCGAGAAGGCGCCGTGCGAGTACGTCTGTCCGGTGAACGCGACGG
TCCACAGCCCCGACGGCCTCAACGAGATGATCTACAACCGATGCATCGGG
ACGCGCTTTTGCTCCAACAACTGCCCGTACAAGATCCGGCGGTTCAATTT
CTTCGACTACAATGCCCACGTCCCGTACAACGCCGGCCTCCGCAAGCTCC
AGCGCAACCCGGACGTGACCGTCCGCGCCCGCGGCGTCATGGAGAAATGC
ACGTACTGCGTGCAGCGGATCCGAGAGGCGGACATCCGCGCGCAGATCGA
GCGGCGGCCGCTCCGGCCGGGCGAGGTGGTCACCGCCTGCCAGCAGGCCT
GTCCGACCGGCGCGATCCAGTTCGGGTCGCTGGATCACGCGGATACCAAG
ATGGTCGCGTGGCGCAGGGAGCCGCGCGCGTACGCCGTGCTCCACGACCT
CGGCACCCGGCCGCGGACGGAGTACCTCGCCAAGATCGAGAACCCGAACC
CCGAGATTGAATGAGCCATGGCGGGCCCGCTCATCCTGGACGCACCGACC
GACGATCAGCTGTCGAAGCAGCTCCTCGAGCCGGTATGGAAGCCGCGCTC
CCGGCTCGGCTGGATGCTCGCGTTCGGGCTCGCGCTCGGCGGCACGGGCC
TGCTCTTCCTCGCGATCACCTACACCGTCCTCACCGGGATCGGCGTGTGG
GGCAACAACATCCCGGTCGCCTGGGCCTTCGCGATCACCAACTTCGTCTG
GTGGATCGGGATCGGCCACGCCGGGACGTTCATCTCCGCGATCCTCCTCC

TGCTCGAGCAGAAGTGGCGGACGAGCATCAACCGCTTCGCCGAGGCGATG
ACGCTCTTCGCGGTCGTCCAGGCCGGCCTCTTTCCGGTCCTCCACCTCGG
CCGCCCCTGGTTCGCCTACTGGATCTTCCCGTACCCCGCGACGATGCAGG
TGTGGCCGCAGTTCCGGAGCGCGCTGCCGTGGGACGCCGCCGCGATCGCG
ACCTACTTCACGGTGTCGCTCCTGTTCTGGTACATGGGCCTCGTCCCGGA
TCTGGCGGCGCTGCGCGATCACGCCCCGGGCCGCGTCCGGCGGGTGATCT
ACGGGCTCATGTCGTTCGGCTGGCACGGCGCCGCCGACCACTTCCGGCAT
TACCGGGTGCTGTACGGGCTGCTCGCGGGGCTCGCGACGCCCCTCGTCGT
CTCGGTGCACTCGATCGTGAGCAGCGATTTCGCGATCGCCCTGGTCCCCG
GCTGGCACTCGACGCTCTTTCCGCCGTTCTTCGTCGCGGGCGCGATCTTC
TCCGGGTTCGCGATGGTGCTCACGCTGCTCATCCCGGTGCGGCGGATCTA
CGGGCTCCATAACGTCGTGACCGCGCGCCACCTCGACGATCTCGCGAAGA
TGACGCTCGTGACCGGCTGGATCGTCATCCTCTCGTACATCATCGAGAAC
TTCCTCGCCTGGTACAGCGGCTCGGCGTACGAAATGCATCAGTTTTTTCA
GACACGCCTGCGCGGCCCGAACAACGCCGCCTACTGGGCCCAGCACGTCT
GCAACGTGCTCGTCATCCAGCTCCTCTGGAGCGAGCGGATCCGGACGAGC
CCCGTCGCGCTCTGGCTCATCTCCATCCTCGTCAACGTCGGGATGTGGAG
CGAGCGGTTCACGCTCATCGTGATGTCGCTCGAGGAAGAGTTCCTCCCGT
CCAAGTGGCACGGCTACAGCCCGACGTGGGTGGACTGGAGCCTCTTCATC
GGGTCAGGCGGCTTCTTCATGCTCCTGTTCCTGAGCTTTTTGCGCGTCTT
TCCGTTCATCCCCGTCGCGGAGGTCAAGGAGCTCAACCATGAAGAGCTGG
AGAAGGCTCGGGGCAAGGGGGGCGCTGATGGAGACCGGAACGCTCGGCG
AGTTCGACGACCCGGAGGCGATGCTCCATGCGATCCGAGAGCTCAGGCGG
CGCGGCTACCGCCGGGTGGAAGCGTTCACGCCCTATCCGGTGAAGGGGCT
CGACGAGGCGCTCGACCTCCCGCGCTCGAACCTCAACCGGATGGTGCTGC
CCTTCGCGATCCTGGGGGTCGTGGGCGGCTACTTCGTCCAGTGGTTCTGC
AACGCTTTCCACTATCCGCTGAACGTGGGCGGGCGCCCGCTGAACTCGGC
GCCGGCGTTCATCCCGATCACGTTCGAGATGGGGGTGCTCTCCACCTCGA
TCTTCGGCGTGCTCATCGGCTTTTACCTGACGAGGCTGCCGAGGCTCTAC
CTCCCGCTCTTCGACGCCCCGGGCTTCGAGCGCGTCACGCTGGATCGGGT

CCTGGTCGGGCTCGACGACACGGAACCTTCCTTCTCGAGCGCCCAGGCGG
AGCGCGATCTCCTCGCGCTCGGCGCCAGGCGCGTCGTCGTGGCGAGGAGG
CGCGAGGAGCCATGAGGGCCGGCGCCCCGGCTCGCCCCCTCGGGCGCGCG
CTCGCGCCGTTCGCCCTCGTCCTGCTCGCCGGGTGCCGCGAGAAGGTGCT
GCCCGAGCCGGACTTCGAGCGGATGATCCGCCAGGAGAAATACGGGCTCT
GGGAGCCGTGCGAGCACTTCGACGACGGCCGCGCGATGCAGCACCCGCCC
GAGGGGACCGTCGCGCGCGGGCGCGTCACCGGGCCGCCCGGCTATCTCCA
GGGCGTCCTCGACGGGGCGTACGTCACGGAGGTGCCGCTCTCGCTCACGG
TCGAGCTCGTGCAGCGCGGCCGGCAGCGCTTCGAGACCTTCTGCGCGCCG
TGCCACGGGATCCTCGGCGACGGCAGCTCGCGCGTGGCGACGAACATGAC
GCTGCGCCCGCCGCCGTCGCTCGTCGGACCCGAGGCGCGGAGCTTCCCGC
CGGGCAGGATCTACCAGGTCATCATCGAGGGCTACGGCCTGATGCCGCGC
TACTCGGACGATCTGCCCGACATCGAAGAGCGCTGGGCCGTCGTCGCCTA
CGTGAAGGCGCTTCAGCTGAGCCGCGGAGTGGCCGCGGGCGCCCTCCCGC
CCGCGCTCCGCGGCCGGGCAGAGCAGGAGCTGCGATGAACAGGGATGCCA
TCTAGTACAAGGGCGGCGCGACGATCGCGGCCTCGCTCGCGATCGCGGCG
CTCGGCGCGGTCGCCGCGATCGTCGGCGGCTTCGTCGATCTCCGCCGGTT
CTTCTTCTCGTACCTCGCCGCGTGGTCGTTCGCGGTATTCCTGTCCGTGG
GCGCGCTCGTCACGCTCCTCACCTGCAACGCCATGCGCGCGGGCTGGCCC
ACGGCGGTGCGCCGCCTCCTCGAGACGATGGTGGCGCCGCTGCCCCTGCT
CGCGGCGCTCTCCGCGCCGATCCTGGTCGGCCTGGACACGCTGTACCCGT
GGATGCACCCCGAGCGGATCGCCGGCGAGCACGCGCGGCGCATCCTCGAG
CACAGAACGCCCTACTTCAATCCAGGCTTCTTCGTCGTGCGCTCGGCGAT
CTACTTCGCGATCTGGATCGCCGTCGCCCTCGTGCTCCGCCGGCGATCGT
TCGCGCAGGACCGTGAGCCGAGGGCCGACGTCAAGGACGCGATGTATGGC
CTGAGCGGCGCCATGCTGCCGGTCGTGGCGATCACGATCGTCTTCTCGTC
GTTCGACTGGCTCATGTCCCTCGACGCGACCTGGTACTCGACGATGTTCC
CGGTCTACGTGTTCGCGAGCGCCTTCGTGACCGCCGTCGGCGCGCTCACG
GTCCTCTCGTATGCCGCGCAGACGTCCGGTTACCTCGCGAGGCTCAACGA
CTCGCACTATTACGCGCTCGGGCGGCTGCTCCTCGCGTTCACGATATTCT

GGGCCTATGCGGCCTATTTCCAGTTCATGTTGATCTGGATCGCGAACAAG
CCCGACGAGGTCGCCTTCTTCCTCGACCGCTGGGAAGGGCCCTGGCGGCC
GACCTCCGTGCTCGTCGTCCTCACGCGGTTCGTCGTCCCGTTCCTGATCC
TGATGTCGTACGCGATCAAGCGGCGCCCGCGCCAGCTCTCGTGGATGGCG
CTCTGGGTCGTCGCCTCCGGCTACATCGACTTTCACTGGCTCGTGGTGCC
GGCGACAGGGCGCCACGGGTTCGCCTATCACTGGCTCGACCTCGCGACCC
TGTGCGTCGTGGGCGGCCTCTCGACCGCGTTCGCCGCGTGGCGGCTGCGA
GGGCGGCCGGTGGTCCCGGTCCACGACCCGCGGCTCGAAGAGGCCTTTGC
GTACCGGAGCATATGATGTTCCGTTTCCGTCACAGCGAGGTTCGCCAGGA
GGAGGACACGCTCCCCTGGGGGCGCGTGATCCTCGCGTTCGCCGTCGTGC
TCGCGATCGGCGGCGCGCTGACGCTCTGGGCCTGGCTCGCGATGCGGGCC
CGCGAGGCGGATCTGCGGCCCTCCCTCGCGTTCCCCGAGAAGGATCTCGG
GCCGCGGCGCGAGGTCAGCATGGTCCAGCAGTCGCTGTTCGACGAGGCGC
GCCTGGGCCAGCAGCTCGTCGAGGCGCAGCGCGCGGAGCTCCGCCGCTTC
GGCGTCGTCGATCGGGAGAGGGGCATCGTGAGCATCCCGATCGACGACGC
GATCGAGCTCATGGTGGCGGAGGGCGCGCGATGAGCCGGGCCGTCGCCGT
GGCCCTCCTGCTGGCAGCCGGCCTCGTGTCGCGCCCGGGCGCCGCGTCCG
AGCCCGTATCGCTTTCGCCCCGCGCTGGGCCCGTCCGCGGGCGAGGCCGC
GCTCTGAAACGACGGCTCCGGCGCGGATGAGCGGCCCGAGGCGACCTCCT
GCAACCCACCGCGCTGCGTACAGGGGTAAATCAACTGCATTCCAGGATAC
GGACCGCGCAGCATAAACCCTCACCGGACAGACTGAATAATGCCGAACCT
TGAACTTCTATGCCCATGCGTGCGGGATCATCATGCCCAACTATTTAAAC
TGGTCCCTCCGGCAAAAGGAACGGACCAGCACCCAGAATAACCCCTGTTT
GGCCAGCGCAAAAAAATGAAAACTCTTCTCCTTGCGGCTAAAATAAACGA
CTCCGGGGAACCGAATGGTATGCAATACACCATGGCAACGCGATTCGTCC
TAACCTTAAACAAACTTCCCAGCGACTCGTCCGTCGAACGCTTCTGACGC
AAGACCATGACCCCACGAGAACCGGGCGGCGGACACACTGCCAGTGAAAC
TCGGCCTAGGCCCGCCCTGCCTTCACGTATTCACCGTGGGGCGGTCCAAA
ACTAAAAACAATACGTGACTCTTCTCAAATATCGTCGGATAAAGGCCAAC
ACGCGTACCTCCCCCTAAAGGGAAGAAACCCCTACCAGGGTGGACCGTAT

CCACCCGTGATCCCTGAAACATCTCCAGCCGCGTACAAATTAGGCTTTGA
CAAAACC

SEQ ID No 80 (>Contig55)

GGGGAAAGAAGTCGGGAGCAGCAGAAGAGCGCAGCGATTGAAAGCACGAC
GCAAGGGCAAAGGACATNAGACGACAACAGAAAGGCGAAGACATGAGAAG
AGGAGAAAGACTCGGAAATGGGCACAAAGCGAGAAAAAACGTAACGCTAT
CTGAGAAAGAACACCATAAAGTCAGACTGGGGTAAAGCCATACACGCAGC
GAAAGCAGACAAGCAAAAGTCATTAACAGATGCAACAAAGAACAAAAGAG
AAAGGAAGGACACATGAAGGAGAAAGGCGGCTTCGAGAACTAAAGACGGG
AGCCAACAATATACCTTTCACAGGGGCGAAGAAGGGCCCAAGGTTCAATC
GAAGACGATTGAATCCAAGGAAGTCCAATCGGAAGAAAAAAGCATATTGA
TAAACCAGACAACGACAGCAGGCCTGAACGTAGGAGCGAGATCGTGAGAC
ATCAGTAGGCAAAACAAGAGCGCTACACCCAGGGGTCGTCAACCTAGAAA
GGCGCGTCCTCAAGCCGGTAGCGGCCGCGCGCGACCAGCCCGATGCGGGC
GCCTTCGCGCGCGAACCGCCGCACCGTCGCGCGCCCCACGCCCGCTGACG
CCCCCGTGATCACCACGACCTCCGGCCGCCTGGAATCGCCCATGCCCGTC
GCCTCCCGCCTCGCCCGCGGCGTCAAGCAACGTGAATGCCACCTGAGCGT
GTCACTTCCTCAAGCTCGACAGCACGTCCTTGATCCGCTCGGTCGGACCC
GCCGTGAGCGACCAGGGGCTCGGCTGCCGCACGACGTGCAGGTCGCCTCG
ACGCTCGAGCACCTCGAACCGCGTCGTTCCCTGCTCGGTCCGGACGAACC
GGAGCGCCACCACGGCCGTGCCCACATGGAGGTTCGAGAGGGTGATCTCG
GGCAGCCACGCCGGCAGCGCCGGGTCGACCAGCAGCAGGTCGAGCGGCGC
GAACGGGTACAGCCCGAGCATCGCCTGGAGCATCGCGAACACGCTCGAGC
AGGACCAGGCCTGCGGCCAGTTGGCCTTCGGGTACAGCGCGGGGAACGGG
TGCTCCGCGTCGCGCGGGTGGCCGCTCCAGCACTCGGGCAGCCGGTGGTG
CTCGAACAGCGCCGCCGCCTCGAACACGGCGCGGCACAGGAGCGCCACGT
GCCCGTGCAGGCCGTACCGCGCGAGCCCGAGCGCGATCGCGCCCTGATCG
ACGGGCCAGACCGTCCCCCGGTGATAGCTGTACGGATCGAACGCGGGGTG
ATCGGCCGAGAGCGTGCGGATACCCCAGCCGGAGAACATGTCCGCCGCGA

ACATGCGGATGGCGGTGGGCTCGGCGAGGGCCTGGTCCACGATCCCCGCC
GCGAGGCAGAGCCCCGGATCCGAGCCGATCGAGCGGATCTGGCGCTTGTC
CGGGCCGAGGCCCATCGCGAAGGTGCGCGCGTCGGGCATCCAGAAGGCGT
CGTTGAACCGCCTCTGCAGCTCGAGCGCCTCGGCGAACAGCCGTCGCGCG
TCGTCCTTGCGGCCGAACCAGAAGAGCAGCTCGGAGAGGCGCAGCTTCGA
CAGGAACACGAAGCCCTGCATCTCGCACGTCCCGATCGGCGGCCGCACCT
GAGAGCCGTCGGCGTGGACGATGGCGTCGTCGGAGTCCTTCCAGCCCTGG
TTCTGGATCGACGCGCTCGAGCGGGGCTCGTACTCGTAGAACCCGTCGCC
GTCGAGATCGCCCTCCTCGTCGATCCAGCGCATGGCCCTGAGCGCAGGCT
CGATCAGGCGGCCGACGCGCTCGCGATCGCCGGTCCAGTGCCAGAGCTCC
GAGGCGGCCACCGCGTAGAACATCGTCGAGGTCGCCGACGCGTACGTGCG
CCCCAGCGGGTTGTAGTTGAGGTCCGAGAGCGCTCCGTCCCTGGCCTGAT
GCAGCATCCGGTCGGGCTGCTCGTCGCGCCAGTCGTCGACGACGCGCCCC
TGCCAGCGCGGGAGCACGAGCGCTGTGCCTGCGAGGATGTCGGTCGTCAG
CGCGGCGGCCTGCGTGCCCGCGGCGAGCGGGTCGCGGCCGAAGAGCCCGA
TGTAGATCGGCAGTCCGGCGGCCACCGTCCAGGAGCGCTCGTCCTGGTCG
ATATCGTACATGCGGAGCGCGATGAGATCGCGCTTGGCTCGCTCGAGCAC
GGAGAAGACCGTATGCGAGAGCGTGTCGGCGCCCGGGACCGAGAAGGACG
TGGCGCGGTGGTGGAAGGTGCCGCGCGCCACGTCGCGCGCGTTTCGCGTG
CCGAAGAACGAGCGGCAGCCGGCGAGCAGCGGGAGCGGTTCGCCGCGGAT
CAGGGCGATCACGTCGACGCAGCCGCGCCAGGTGCCATGGGGCTCGAGAT
CGATCGAGAACCGGATCTCGCGGCCAGCGCAGGACGGGGGCGAGCCCGCG
CTCCGGGCCCGTACGACGATCCCGGCGTCGAAGCGCGCGACGCCGGATTC
TCCGGGGTGTTCATAGCGGTGCTCGGCGCGGTAGTCGCAGCGCAGCGCGC
ATCCGTGCTCGGCGGGCTCGAGCGCCCACCGGGCGTCGCCGCGCTGGAGG
CGCGGGCCGTCCGTCTCCTCCGCGTCGGCGAAGTCGGCGTCGATCTCGAG
CGCCAGCGTGAAGCGCACGCGCTCCTGCGTGAAGCTCGCGACGTCGATGT
CCTCGTGGAAGCCGTCCCCCACGAAGCGCGACAGCCGCAGCTCGACGGTC
CGCGGCGCCGCGCCGCCGCCCTCGGGCGGCGGGGCGATGTAGTAGCCGAG
CCAGCTGTCCGGCTCGACGGCGGAGAGCGCGACGGGCGCCGGCGGCCTGC

CGTCGATCAGGTGGCGGTAGAGGGAGAGGAGGCGGGTGTTGCGCACGAAC
AGGCCGATGTGCGGCTCGGGGGCGATCGAGCCGTCCGGACGCATGCACAG
GACCGTCCGGTTCTGGCTGATGTACAGGGAGCCGGCGCGGGGCCTCAGCG
TGGCCAGCGAGCCGAAGGGTCTCTCGAGCGACATGGTCACCTCCAGGCGA
GCCCGGGGTGAGGACTGCCAGGGGCGTGCCAACGACGTGGACGCGCTCAC
GCCAAGCCGCTGGGCGGGCGCGGCGGCCCCTGCGCCCGAGGCTCAGCCGA
GCGCGAGCGGTGTTTCGTCGGCGCGAGCGGGCCTCGTCGTCGCCTCCAGG
TAGGCCTTTCCGATGTCGGCCTCGTCCACGAACCCGACGATCGCGCCCTC
GCCGTCGACGACGGGGACCTCGCGGACGCCGTGCGCCACCATCGCCTCGG
TCGCCGTCCGCAGATCGTCGGTGACCGTCACGGCCACCGGCGGCTGCATC
GCGTCGGCGGCCACGGTCATCCGCTCGAGGTCGTGCTCCACCGCGATGAT
CCGGAGCGACTCGGCGGTGATCATGCCGACCATCTTGCGCGACGGTTCGA
GCACCGGGAACACCTCCTGCCAGCTCGCGTCGGCCGCCCGCCGGAGCATC
TCGCGGGCCGGCGTCCCCGGCACGAACGTCACGAGCGCGCGCCCCTCGAT
CATGATCTGCCGCACGCGGATGGTCTTGAGCACGTCGAGCGTCGGCACCG
GGTGCGCGGGAGACTCGCGCTGGGTGGGGAGCTGCGCGTGGTAGAGCGAG
TGCTTCCGCAGCGCGACGAAGGCGACGCCCTCGGCGAGCATCAGCGGAAC
CAGGAGGTCGTAGCTGCCGGCGAGCTCGCAGACCATCACGAGGGAGCTCA
CCGGCACGTGCGCGACGCCGCCGTAGAAGGTGCCCATGCCCACGAGCGCG
AAGGCGCCCGGGTCGATGCGCGGATCGCCGAGCAGGAGCGCCGCCGCGCG
CCCGAACGCGCCGCCGAAGAGCCCGCCGATGACGAGCGACGGCGCGAAGT
CGCCGGCGCACCCGCCGCTGCCGAGCGTGAGCGACGAGGCGACGATCTTG
GCGGCGCAGAGCAGGAGCAGGAGCTCCACGCCGCGCCAGCCCGGGTGGAG
CCACGTGGCGCCGGTGATCGCGACCTGGACGGCGCCGTACCCGCCGCCGA
GCAGCCCGAGCCCTTGCCCGGGGCTCTCGATCCTGCGGCCGACGAACCAG
AGGACGGGCACGCAGAAGAGGCCCAGCGCGAGCCCGCCGAGCCCCGGGCG
CGCCCAGGGGGCGATGGGCAGGCGCGCCGCGATCGCCTTCACGCCGCCGA
GGCACTTCAGGAAGCCGATCGCGACGAAGGCCAGGAGCAGCGCGAGCAGC
GCATAGAGAGGGAGGTGCGACGGGACGAACGCATACTTCGGCGCGTGCGC
GAAGAGCGTCGACTCGCCGTAGAACGAGATGAAGACCGAGTAGGAGACCA

CGCTGGCGAGCAGCGCCGGGATCAGCGCCTCGGCCTCGAAGTCGTCGCGG
TAGAGCACCTCGACGGCGAGCAGGGCGGCGCCGAGCGGCGTGCGGAAGAT
GGCGGACATCCCGGCCGCGACCCCCGCGAGCATCAGGATGCGGTGCTCGC
GCCGGCCGACCGCGAGCCCGCGCCCCACGAGCGAGCCGAGCGCTCCTCCG
ACCTGCATGGTCGGCCCCTCGCGACCGCCGGCGCCGCCCGAGCCGAGCGT
GAGGATCGACGCGACCGCCTTGACCCACGCGACCCGCTTGCGCATCCGGC
CGCCGTGGTGGTGGAAGGCCTGGATCATCGCGTCGCCGCCGCCGCCCGCG
GCCTCGGGGGCGAGGCGCCAGGTGAGGATGCCCCCGGCCAGCGCTCCGAG
CGCCGGGATCAGCAGCAGCAGCCAGAGCCGGACGCTCCGGTGCTCGGGGC
CGTCGCCGCTGAAGATGGCCTCGCCGTGGGCTCGAAGCCGCGCGTAGCCG
GCGAGGCGGCCGAGGAGCAGCTCCTCGACGAGCTCGAGGGCGCCGAAGAA
GAGCACCGCGACGAGGCCCGCGATGGCCCCGACGAGCACGGCATGCAGGA
TCGTGCGCCCGACGAGCCTGAGATCGAGGGGCGCGACCTCCGAGAGGAGC
GCGGAGAAGGGGCGCCGCCGTCGCACCACGCCGGCGCGCTCTGCTGGGAA
TGGTTCCGTCACGGCGATGGT

SEQ ID No 81 (>Contig56)

GGATCCGGCCGCGAGCGGCTGCGGTGCCGATGCGCTCGACGGTGACGGGC
GGGGTGATCGCGGGTCCGGAGCTCGGTGCGAGCTACTGGGCGGACAACCT
TCGGCAGCCGGTGCGCTTCGCTGCGGCGGCGCAAGCGCTGCTGGAGGGTG
GCCCCGCGCTGTTCATCGAGATGAGCCCGCACCCGATCCTGGTGCCGCCC
CTGGACGAGATCCAGACGGCGGCCGAGCAAGGGGGCGCTGCGGTGGGCTC
GCTGCGGCGAGGGCAGGACGAGCGCGCGACGCTGCTGGAGGCGCTGGGGA
CGCTGTGGGCGTCCGGCTATCCGGTGAGCTGGGCTCGGCTGTTCCCCGCG
GGCGGCAGGCGGGTTCCGCTGCCGACCTATCCCTGGCAGCACGAGCGGTG
CTGGATCGAGGTCGAGCCTGACGCCCGCCGCCTCGCCGCAGCCGACCCCA
CCAAGGACTGGTTCTACCGAACGGACTGGCCCGAGGTGCCCCGCGCCGCC
CCGAAATCGGAGACAGCTCATGGCAGCTGGCTGCTGTTGGCCGACAGGGG
TGGGGTCGGTGAGGCGGTCGCTGCAGCGCTGTCGACGCGCGGACTTTCCT
GCACCGTGCTTCATGCGTCGGCTGACGCCTCCACCGTCGCCGAGCAGGTA

TCCGAAGCTGCCAGTCGCCGAAACGACTGGCAGGGAGTCCTCTACCTGTG
GGGCCTCGACGCCGTCGTCGATGCTGGGGCATCGGCCGACGAAGTCAGCG
AGGCTACCCGCCGTGCCACCGCACCCGTCCTTGGGCTGGTTCGATTCCTG
AGCGCTGCGCCCCATCCTCCTCGCTTCTGGGTGGTGACCCGCGGGGCATG
CACGGTGGGCGGCGAGCCAGAGGCCTCTCTTTGCCAAGCGGCGTTGTGGG
GCCTCGCGCGCGTCGCGGCGCTGGAGCACCCCGCTGCCTGGGGTGGCCTC
GTGGACCTGGATCCTCAGAAGAGCCCGACGGAGATCGAGCCCCTGGTGGC
CGAGCTGCTTTCGCCGGACGCCGAGGATCAACTGGCGTTCCGCAGCGGTC
GCAGGCACGCAGCACGCCTTGTAGCCGCCCCGCCGGAGGGCGACGTCGCA
CCGATATCGCTGTCCGCGGAGGGGAGCTACCTGGTGACGGGCGGCTGGG
TGGCCTTGGTCTGCTCGTGGCTCGGTGGCTGGTGGAGCGGGGAGCTCGAC
ATCTGGTGCTCACCAGCCGGCACGGGCTGCCAGAGCGACAGGCGTCGGGC
GGAGAGCAGCCGCCGGAGGCCCGCGCGCATCGCAGCGGTCGAGGGGCT
GGAAGCGCAGGGCGCGCGGGTGACCGTGGCAGCGGTGGATGTCGCCGAGG
CCGATCCCATGACGGCGCTGCTGGCCGCCATCGAGCCCCCGTTGCGCGGG
GTGGTGCACGCCGCCGGCGTCTTCCCCGTGCGTCACCTGGCGGAGACGGA
CGAGGCCCTGCTGGAGTCGGTGCTCCGTCCCAAGGTGGCCGGGAGCTGGC
TGCTGCACCGGCTGCTGCGCGACCGGCCTCTCGACCTGTTCGTGCTGTTC
TCGTCGGGCGCGGCGGTGTGGGGTGGCAAAGGCCAAGGCGCATACGCCGC
GGCCAATGCGTTCCTCGACGGGCTCGCGCACCATCGCCGCGCGCACTCGC
TGCCGGCGTTGAGCCTCGCCTGGGGCTTATGGGCCGAGGGAGGCATGGTT
GATGCAAAGGCTCATGCACGTCTGAGCGACATCGGGGTCCTGCCCATGGC
CACGGGGCCGGCCTTGTCGGCGCTGGAGCGCCTGGTGAACACCAGCGCTG
TCCAGCGTTCGGTCACACGGATGGACTGGGCGCGCTTCGCGCCGGTCTAT
GCCGCGCGAGGGCGGCGCAACTTGCTTTCGGCTCTGGTCGCGGAGGACGA
GCGCGCTGCGTCTCCCCCGGTGCCGACGGCAAACCGGATCTGGCGCGGCC
TGTCCGTTGCGGAGAGCCGCTCAGCCCTCTACGAGCTCGTTCGCGGCATC
GTCGCCCGGGTGCTGGGCTTCTCCGACCCGGGCGCGCTCGACGTCGGCCG
AGGCTTCGCCGAGCAGGGGCTCGACTCCCTGATGGCTCTGGAGATCCGTA
ACCGCCTTCAGCGCGAGCTGGGCGAACGGCTGTCGGCGACTCTGGCCTTC

GACCACCCGACGGTGGAGCGGCTGGTGGCGCATCTCCTCACCGACGTGCT
GAAGCTGGAGGACCGGAGCGACACCCGGCACATCCGGTCGGTGGCGGCGG
ATGACGACATCGCCATCGTCGGTGCCGCCTGCCGGTTCCCAGGTGGGGAT
GAGGGCCTGGAGACATACTGGCGGCATCTGGCCGAGGGCATGGTGGTCAG
CACCGAGGTGCCAGCCGACCGGTGGCGCGCGGCGGACTGGTACGACCCCG
ATCCGGAGGTTCCGGGCCGGACCTATGTGGCCAAGGGTGCCTTCCTCCGC
GATGTGCGCAGCTTGGATGCGGCGTTCTTCGCCATTTCCCCTCGTGAGGC
GATGAGCCTGGACCCGCAACAGCGGCTGTTGCTGGAGGTGAGCTGGGAGG
CGATCGAGCGCGCTGGCCAGGACCCGATGGCGCTGCGCGAGAGCGCCACG
GGCGTGTTCGTGGGCATGATCGGGAGCGAGCACGCCGAGCGGGTGCAGGG
CCTCGACGACGACGCGGCGTTGCTGTACGGCACCACCGGCAACCTGCTCA
GCGTCGCCGCTGGACGGCTGTCGTTCTTCCTGGGTCTGCACGGCCCGACG
ATGACGGTGGACACCGCCTGCTCGTCGTCGCTGGTGGCGTTGCACCTCGC
CTGCCAGAGCCTGCGATTGGGCGAGTGCGACCAGGCCCTGGCCGGCGGGT
CCAGCGTGCTTTTGTCGCCGCGGTCATTCGTCGCGGCGTCGCGCATGCGT
TTGCTTTCGCCAGATGGGCGGTGCAAGACGTTCTCGGCCGCTGCAGACGG
CTTTGCGCGGGCCGAGGGCTGCGCCGTGGTGGTGCTCAAGCGGCTCCGTG
ACGCGCAGCGCGACCGCGACCCCATCCTGGCGGTGGTCAGGAGCACGGCG
ATCAACCACGATGGCCCGAGCAGCGGGCTCACGGTGCCCAGCGGTCCTGC
CCAGCAGGCGTTGCTACGCCAGGCGCTGGCGCAAGCGGGCGTGGCGCCGG
CCGAGGTCGATTTCGTGGAGTGCCACGGGACGGGGACAGCGCTGGGTGAC
CCGATCGAGGTGCAGGCGCTGGGCGCGGTGTACGGGCGGGGCCGCCCCGC
GGAGCGGCCGCTCTGGCTGGGCGCTGTCAAGGCCAACCTCGGCCACCTGG
AGGCCGCGGCGGGCTTGGCCGGCGTGCTCAAGGTGCTCTTGGCGCTGGAG
CACGAGCAGATTCCGGCTCAACCGGAGCTCGACGAGCTCAACCCGCACAT
CCCGTGGGCAGAGCTGCCAGTGGCCGTTGTCCGCAGGGCGGTCCCCTGGC
CGCGCGGCGCGCGCCCGCGTCGTGCAGGCGTGAGCGCTTTCGGCCTGAGC
GGGACCAACGCGCATGTGGTGTTGGAGGAGGCGCCGGCGGTGGAGCCTGT
GGCCGCGGCCCCCGAGCGCGCAGCGGAGCTGTTCGTCCTGTCGGCGAAGA
GCGCGGCGGCGCTGGATGCGCAGGCAGCCCGGCTGCGGGACCACCTGGAG

AAGCATGTCGAGCTTGGCCTCGGCGATGTGGCGTTCAGCCTGGCGACGAC
GCGCAGCGCGATGGAGCACCGGCTGGCGGTGGCCGCGAGCTCGCGCGAGG
CGCTGCGAGGGGCGCTTTCGGCCGCAGCGCAGGGGCACACGCCGCCGGGA
GCCGTGCGTGGGCGGGCCTCGGGCGGCAGCGCGCCGAAGGTGGTCTTCGT
GTTTCCCGGCCAGGGCTCGCAGTGGGTGGGCATGGGCCGAAAGCTCATGG
CCGAAGAGCCGGTCTTCCGGGCGGCGCTGGAGGGTTGCGACCGGGCCATC
GAGGCGGAAGCGGGCTGGTCGCTGCTCGGGGAGCTCTCCGCCGACGAGGC
CGCCTCGCAGCTCGGGCGCATCGACGTGGTTCAGCCGGTGCTGTTCGCCA
TGGAAGTAGCGCTTTCTGCGCTGTGGCGGTCGTGGGGAGTGGAGCCGGAA
GCGGTGGTGGGCCACAGCATGGGCGAGGTTGCGGCGGCGCACGTGGCCGG
CGCGCTGTCGCTCGAGGACGCGGTGGCGATCATCTGCCGGCGCAGCCGGC
TGCTGCGGCGGATCAGCGGTCAGGGGGAGATGGCGCTGGTCGAGCTGTCG
CTGGAGGAGGCCGAGGCGGCGCTGCGTGGCCATGAGGGTCGGCTGAGCGT
GGCGGTGAGCAACAGCCCGCGCTCGACCGTGCTCGCCGGCGAGCCGGCGG
CGCTCTCGGAGGTGCTGGCGGCGCTGACGGCCAAGGGGGTGTTCTGGCGG
CAGGTGAAGGTGGACGTCGCCAGCCATAGCCCGCAGGTCGACCCGCTGCG
CGAAGAGCTGATCGCGGCGCTGGGAGCGATCCGGCCGCGAGCGGCTGCGG
TGCCGATGCGCTCGACGGTGACGGGCGGGGTGATCGCGGGTCCGGAGCTC
GGTGCGAGCTACTGGGCGGACAACCTTCGGCAGCCGGTGCGCTTCGCTGC
GGCGGCGCAAGCGCTGCTGGAGGGTGGCCCCGCGCTGTTCATCGAGATGA
GCCCGCACCCGATCCTGGTGCCGCCCCTGGACGAGATCCAGACGGCGGCC
GAGCAAGGGGGCGCTGCGGTGGGCTCGCTGCGGCGAGGGCAGGACGAGCG
CGCGACGCTGCTGGAGGCGCTGGGGACGCTGTGGGCGTCCGGCTATCCGG
TGAGCTGGGCTCGGCTGTTCCCCGCGGGCGGCAGGCGGGTTCCGCTGCCG
ACCTATCCCTGGCAGCACGAGCGGTACTGGATCGAGGACAGCGTGCATGG
GTCGAAGCCCTCGCTGCGGCTTCGGCAGCTTCGCAACGGCGCCACGGACC
ATCCGCTGCTCGGGGCTCCATTGCTCGTCTCGGCGCGACCCGGAGCTCAC
TTGTGGGAGCAAGCGCTGAGCGACGAGAGGCTATCCTACCTTTCGGAACA
TAGGGTCCATGGCGAAGCCGTGTTGCCCAGCGCGGCGTATGTAGAGATGG
CGCTCGCCGCCGGCGTAGATCTCTATGGCACGGCGACGCTGGTGCTGGAG

CAGCTGGCGCTCGAGCGAGCCCTCGCCGTGCCCTCCGAAGGCGGACGCAT
CGTGCAAGTGGCCCTCAGCGAAGAAGGTCCCGGTCGGGCCTCATTCCAGG
TATCGAGTCGTGAGGAGGCAGGTAGGAGCTGGGTGCGGCACGCCACGGGG
CACGTGTGTAGCGGCCAGAGCTCAGCGGTGGGAGCGTTGAAGGAAGCTCC
GTGGGAGATTCAACGGCGATGTCCGAGCGTCCTGTCGTCGGAGGCGCTCT
ATCCGCTGCTCAACGAGCACGCCCTCGACTATGGTCCCTGCTTCCAGGGC
GTGGAGCAGGTGTGGCTCGGCACGGGGGAGGTGCTCGGCCGGGTACGCTT
GCCAGGAGACATGGCATCCTCAAGTGGCGCCTACCGGATTCATCCCGCCT
TGTTGGATGCATGTTTTCAGGTGCTGACAGCGCTGCTCACCACGCCGGAA
TCCATCGAGATTCGGAGGCGGCTGACGGATCTCCACGAACCGGATCTCCC
GCGGTCCAGGGCTCCGGTGAATCAAGCGGTGAGTGACACCTGGCTGTGGG
ACGCCGCGCTGGACGGTGGACGGCGCCAGAGCGCGAGCGTGCCCGTCGAC
CTGGTGCTCGGCAGCTTCCATGCGAAGTGGGAGGTCATGGAGCGCCTCGC
GCAGGCGTACATCATCGGCACTCTCCGCATATGGAACGTCTTCTGCGCTG
CTGGAGAGCGTCACACGATAGACGAGTTGCTCGTCAGGCTTCAAATCTCT
GTCGTCTACAGGAAGGTCATCAAGCGATGGATGGAACACCTTGTCGCGAT
CGGCATCCTTGTAGGGGACGGAGAGCATTTTGTGAGCTCTCAGCCGCTGC
CGGAGCCTGATTTGGCGGCGGTGCTCGAGGAGGCCGGGAGGGTGTTCGCC
GACCTCCCAGTCCTATTTGAGTGGTGCAAGTTTGCCGGGGAACGGCTCGC
GGACGTATTGACCGGTAAGACGCTCGCGCTCGAGATCCTCTTCCCTGGTG
GCTCGTTCGATATGGCGGAGCGAATCTATCGAGATTCGCCCATCGCCCGT
TACTCGAACGGCATCGTGCGCGGTGTCGTCGAGTCGGCGGCGCGGGTGGT
AGCACCGTCGGGAATGTTCAGCATCTTGGAGATCGGAGCAGGGACGGGCG
CGACCACCGCCGCCGTCCTCCCGGTGTTGCTGCCTGACCGGACGGAGTAC
CATTTCACCGATGTTTCTCCGCTCTTCCTTGCTCGCGCGGAGCAAAGATT
TCGAGATTATCCATTCCTGAAGTATGGCATTCTGGATGTCGACCAGGAGC
CAGCTGGCCAGGGATACGCACATCAGAGGTTTGACGTCATCGTCGCGGCC
AATGTCATCCATGCGACCCGCGATATAAGAGCCACGGCGAAGCGTCTCCT
GTCGTTGCTCGCGCCCGGAGGCCTTCTGGTGCTGGTCGAGGGCACAGGGC
ATCCGATCTGGTTCGATATCACCACGGGATTGATTGAGGGGTGGCAGAAG

```
TACGAAGATGATCTTCGTATCGACCATCCGCTCCTGCCTGCTCGGACCTG
GTGTGACGTCCTGCGCCGGGTAGGCTTTGCGGACGCCGTGAGTCTGCCAG
GCGACGGATCTCCGGCGGGGATCCTCGGACAGCACGTGATCCTCTCGCGC
GCGCCGGGCATAGCAGGAGCCGCTTGTGACAGCTCCGGTGAGTCGGCGAC
CGAATCGCCGGCCGCGCGTGCAGTACGGCAGGAATGGGCCGATGGCTCCG
CTGACGTCGTCCATCGGATGGCGTTGGAGAGGATGTACTTCCACCGCCGG
CCGGGCCGGCAGGTTTGGGTCCACGGTCGATTGCGTACCGGTGGAGGCGC
GTTCACGAAGGCGCTCGCTGGAGATCTGCTCCTGTTCGAAGACACCGGGC
AGGTCGTGGCAGAGGTTCAGGGGCTCCGCCTGCCGCAGCTCGAGGCTTCT
GCTTTCGCGCCGCGGGACCCGCGGGAAGAGTGGTTGTACGCTTTGGAATG
GCAGCGCAAAGACCCTATACCAGAGGCTCCGGCAGCCGCGTCTTCTTCCT
CCGCGGGGGCTTGGCTCGTGCTGATGGACCAGGGCGGACAGGCGCTGCG
CTCGTATCGCTGCTGGAAGGGCGAGGCGAGGCGTGCGTGCGCGTCATCGC
GGGTACGGCATACGCCTGCCTCGCGCCGGGGCTGTATCAAGTCGATCCGG
CGCAGCCAGATGGCTTTCATACCCTGCTCCGCGATGCATTCGGCGAGGAC
CGGATTTGTCGCGCGGTAGTGCATATGTGGAGCCTTGATGCGACGGCAGC
AGGGGAGAGGGCGACAGCGGAGTCGCTTCAGGCCGATCAACTCCTGGGGA
GCCTGAGCGCGCTTTCTCTGGTGCAGGCGCTGGTGCGCCGGAGGTGGCGC
AACATGCCGCGGCTTTGGCTCTTGACCCGCGCCGTGCATGCGGTGGGCGC
GGAGGACGCAGCGGCCTCGGTGGCGCAGGCGCCGGTGTGGGGCCTCGGTC
GGACGCTCGCGCTCGAGCATCCAGAGCTGCGGTGCACGCTCGTGGACGTG
AACCCGGCGCCGTCTCCAGAGGACGCAGCCGCACTGGCGGTGGAGCTCGG
GGCGAGCGACAGAGAGGACCAGGTCGCATTGCGCTCGGATGGCCGCTACG
TGGCGCGCCTCGTGCGGAGCTCCTTTTCCGGCAAGCCTGCTACGGATTGC
GGCATCCGGGCGGACGGCAGCTATGTGATCACCGATGGCATGGGGAGAGT
GGGGCTCTCGGTCGCGCAATGGATGGTGATGCAGGGGGCCCGCCATGTGG
TGCTCGTGGATCGCGGCGGCGCTTCCGAGGCATCCCGGGATGCCCTCCGG
TCCATGGCCGAGGCTGGCGCGGAGGTGCAGATCGTGGAGGCCGACGTGGC
TCGGCGCGACGATGTCGCTCGGCTCCTCTCGAAGATCGAACCGTCGATGC
CGCCGCTTCGGGGGATCGTGTACGTGGACGGGACCTTCCAGGGCGACTCC
```

TCGATGCTGGAGCTGGATGCCCGTCGCTTCAAGGAGTGGATGTATCCCAA
GGTGCTCGGAGCGTGGAACCTGCACGCGCTGACCAGGGATAGATCGCTGG
ACTTCTTCGTCCTGTATTCCTCGGGCACCTCGCTTCTGGGCTTGCCAGGA
CAGGGGAGCCGCGCCGCCGGTGACGCCTTCTTGGACGCCATCGCGCATCA
CCGGTGCAAGGTGGGCCTTACAGCGATGAGCATCAACTGGGGATTGCTCT
CCGAAGCATCATCGCCGGCGACCCCGAACGACGGCGGAGCACGGCTCGAA
TACCGGGGGATGGAAGGCCTCACGCTGGAGCAGGGAGCGGCGGCGCTCGG
GCGCTTGCTCGCACGACCCAGGGCGCAGGTAGGGGTGATGCGGCTGAATC
TGCGCCAGTGGTTGGAGTTXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXTATGGTATAACTTATTGATTATAATACAGTATACAAAGGTCCCTTT
TCAGGGACCCTTTCGTATGTTGTAGCTGATTTTATTTTCTTCTTTTCTTT
TGGGTGCTTTTAATAGCCATATATGATGTTTCATATAAAGTTAAAAGTTT
CCCCATCTGTCTATCTATAGCGTCATGTTTTTCAGGATTTCTTAATTTCT
GCAGCAGTTTAGATATATATTTCGCAGGTAACTTTATTATTAATGAATTA
TGGGATTGATACATTTCCTTTGTCCAATCTAAATCTTTTCCTGATTTCAA
TACCTCACTCTCATCTCTATTAATTATAAAACCGAGCTTTCCATACGGAC
CTGTTAAATATGATTGTAATTGTCTATACTCAGATGGACCTAATTCTTCA
AAATTTTTAGCATCAAAAACAACTTGTCTTGTTTTATAGTCCTCCAATAC
TCGTTTCCAAAAATCAGATTTGCCACCATTGGTGCCTATAATATCTCGTC
TCTGAACTGCGTTACCATTTGGATGGGACTTGATGTCTGTTAGGTGGGAT
GCAAATACTATTCTTAGTGCGTCTAAACACCATTGTTCAAATTCAGTGGC
ACCTTCATTTCCTATCGGTATCTGATCTAGATGAGTAGTTATTTGACCTA
TTGTTTTATTTCTAATGGCTGAATTATCTGAAATAATATTAATATCATAT
TCGTCATTTATCTCCTCTGCTTCCTCTGGAGCAAGTGCATTACGATTTAG
ATTCAAACCAAGCCAGTAACATGGATGAATTAATAATTTCTCATTACTTT
CAAAACCTTTATCTGGAGTTCGCCCGTCATGGCAAAATGAATAGGATGAG
GTGTTTTTATCACGAATACCAACAAATCCTACACTATACAAACTTTGGAG
AATTCCACTTGCCTTTAACAACTGTATTTCAGACGTTATTTTAGGATCTC
CATTTTCTTCGATTAATTCGAAAGATGCTTCTATTTTTTTTAAGCACGTA
TAAACTGTTAATTCAGGTTCAATGCTACGAAATGCACTAGTTATAACCTG

TATTGAAGGAAAGATCTTCTGATACTCTTTCCAGAGATCTTCAAGTCTGG
CCATGGAAATTGACTTGGCTGCATATTCTAGGTCAGTGTTTATGATAGTT
TCTCTATTCTCTCTGAATGCGGAAAAAAAAGCTTCATTCAACAATGATAG
TAAATCCCTGGGCCGGTAAAGGGTAAATTGCAAACATCGCTTAAAACCAT
TCCTCCCTTTAAGATCATCCGCTGTGCATCTATCCCAAACTCGTTGATCT
TTCTCAATATCTAGCTTAAATGCTACTTTCATTCTTTTAGCTGACAGCAT
TAGGAGTTGTGCCCAGTCCCAATGCAACCTTATGACTTGACCCTCTATAT
TTCTCGAGTAATCAGGATCTTCCTTTGATAGCGACCTAAATATATTATCC
CTTAAAAAAATTATTGGACGAATGCATTTTGCTTTTTGATTTAATTCAAT
AGATGCATATGCTAGACCTGCAATGATTCCAATTCCTATATTATCCGGTT
CATACGCCTCATCTAGCTTATCCATTAATATGACAACTTTCCTGTCTGAG
CGTTCAAGAAGTGATACTATATTATTTTCTATTTCTGAGATATTCAAATT
GAATTGAAGATCACCAATTGATTCTTCTGGGTTATTTTCATCTAAATACT
CCTTTGCGACAAGCCTACACTTTCTTAAAATGTCACCTTGTGCAGAATTC
CATTTTTTCAAATGTTCATTCAACAATGTTTCTGATGATATTTGAGATGA
CAATTTGTAATGAGATGATATATATGATGCTATCTCCATTAGCATAGCGT
ATCGCCATAATAGTCTTGTTGCTGCCCTTGCTAAATTAAATGATCCTGTA
AATGGTTTCAACATTGATCTGAAACCAATAATTTGAGAATCGTCTGGTGA
GAAACTCAGGATTAATATTTTTTTGTCTTTCTTCCAATGCTCATTTAGCT
GAATAAATAAAGCACTTTTACCTGTCCCTCGTCTACCAACAACAATGGTC
CTGTCATCAGTTTCAATTAGAGTCCTAAAGTCAGCAGTTTCAATGAAAGC
ATTACTCAACATCTTTTTATCATTTTCTGCTGTCGTATCACCAAACGGAT
TAGACTTTGAAGTAATATTCAATTCCATATTCAACCTTTTATGTTAGTTG
CTTTCATTTATTACTTTATATACTGTTGAACGAGCAATATTCATTGTTTT
TGATATATGTGAGGCACCTAACCCCTGTTGCCACATATTTAATACTGCAT
CTCTATCTATTTTTCTTTTTCTACCAAAAACAACTCCTTTTGCCA

SEQ ID No 82 (>Contig57)
TCATCTATTGTATAGTTTGTATATTGATATGATATAATTATAACATATAA
ACAGTAAACTTTCTCTACGTAGATCGAGGAGAAGACTCAATTTGTTGACA

TCXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXACTACTCGCATACCGTTG
CGCAACAGCGGCGCGAGGAGCAGGACGCATACGACATCACCGGCAATACG
CTCAGCGTCGCCGACGGACGGTTGTCTTATACGCTAGGGCTGCAGGGACC
CTGCCTGACCGTCGACACGGTCTGCTCGTCGTCGCTCGTGGCCATCCACC
TTGCCTGCCGCAGCCTGCGCGCTCGCGAGAGCGATCTCGCGCTGGCGGGA
GGCGTCAACATGCTCCTTTCGTCCAAGACGATGATAATGCTGGGGCGCAT
CCAGGCGCTGTCGCCCGATGGCCACTGCCGGACATTCGACGCCTCGGCCA
ACGGGTTCGTCCGTGGGGAGGGCTGCGGTATGGTCGTGCTCAAACGGCTC
TCCGACGCCCAGCGACACGGCGATCGGATCTGGGCTCTGATCCGGGGTTC
GGCCATGAATCAGGATGGCCGGTCGACAGGGTTGATGGCACCCAATGTGC
TCGCTCAGGAGGCGCTCTTGCGCGAGGCGCTGCAGAGCGCTCGCGTCGAC
GCCGGGGCCATCGGTTATGTCGAGACCCACGGAACGGGGACCTCGCTCGG
CGACCCGATCGAGGTCGAGGCGCTGCGTGCCGTGTTGGGGCCGGCGCGGG
CCGATGGGAGCCGCTGCGTGCTGGGCGCAGTGAAGACAAACCTCGGCCAC
CTGGAGGGCGCTGCAGGCGTGGCGGGTTTGATCAAGGCGGCGCTGGCTCT
GCACCACGAACTGATCCCGCGAAACCTCCATTTCCACACGCTCAATCCGC
GGATCCGGATCGAGGGGACCGCGCTCGCGCTGGCGACGGAGCCGGTGCCG
TGGCCGCGGGCGGGCCGACCGCGCTTCGCGGGGGTGAGCGCGTTCGGCCT
CAGCGGCACCAACGTCCATGTCGTGCTGGAGGAGGCGCCGGCCACGGTGC
TCGCACCGGCGACGCCGGGGCGCTCAGCGGAGCTTTTGGTGCTGTCGGCG
AAGAGCGCCGCCGCGCTGGACGCACAGGCGGCGCGGCTCTCAGCGCACAT
CGCCGCGTACCCGGAGCAGGGTCTCGGAGACGTCGCGTTCAGCCTGGTAT
CGACGCGTAGCCCGATGGAGCACCGGCTCGCGGTGGCGGCGACCTCGCGC
GAGGCGCTGCGAAGCGCGCTGGAGGTTGCGGCGCAGGGGCAGACCCCGGC
AGGCGCGGCGCGCGGCAGGGCCGCTTCCTCGCCCGGCAAGCTCGCCTTCC
TGTTCGCCGGGCAGGGCGCGCAGGTGCCGGGCATGGGCCGTGGGTTGTGG
GAGGCGTGGCCGGCGTTCCGCGAGACCTTCGACCGGTGCGTCACGCTCTT
CGACCGGGAGCTCCATCAGCCGCTCTGCGAGGTGATGTGGGCCGAGCCGG
GCAGCAGCAGGTCGTCGTTGCTGGACCAGACGGCGTTCACCCAGCCGGCG
CTCTTTGCGCTGGAGTACGCGCTGGCCGCGCTCTTCCGGTCGTGGGGCGT

GGAGCCGGAGCTCGTCGCTGGCCATAGCCTCGGCGAGCTGGTGGCCGCCT
GCGTGGCGGGTGTGTTCTCCCTCGAGGACGCCGTGCGCTTGGTGGTCGCG
CGCGGCCGGTTGATGCAGGCGCTGCCGGCCGGCGGCGCGATGGTATCGAT
CGCCGCGCCGGAGGCCGACGTGGCTGCCGCGGTGGCGCCGCACGCAGCGT
TGGTGTCGATCGCGGCAGTCAATGGGCCGGAGCAGGTGGTGATCGCGGGC
GCCGAGAAATTCGTGCAGCAGATCGCGGCGGCGTTCGCGGCGCGGGGGGC
GCGAACCAAACCGCTGCATGTCTCGCACGCGTTCCACTCGCCGCTCATGG
ATCCGATGCTGGAGGCGTTCCGGCGGGTGACTGAGTCGGTGACGTACCGG
CGGCCTTCGATCGCGCTGGTGAGCAACCTGAGCGGGAAGCCCTGCACCGA
TGAGGTGAGCGCGCCGGGTTACTGGGTGCGTCACGCGCGAGAGGCGGTGC
GCTTCGCGGACGGAGTGAAGGCGCTGCACGCGGCCGGTGCGGGCCTCTTC
GTCGAGGTGGGGCCGAAGCCGACGCTGCTCGGCCTTGTGCCGGCCTGCCT
GCCGGATGCCAGGCCGGTGCTGCTCCCAGCGTCGCGCGCCGGGCGTGACG
AGGCTGCGAGCGCGCTAGAGGCGCTGGGTGGGTTCTGGGTCGTCGGTGGA
TCGGTCACCTGGTCGGGTGTCTTCCCTTCGGGCGGACGGCGGGTACCGCT
GCCAACCTATCCCTGGCAGCGCGAGCGTTACTGGATCGAAGCGCCGGTCG
ATCGTGAGGCGGACGGCACCGGCCGTGCTCGGGCGGGGGGCCACCCCCTT
CTGGGTGAAGTCTTTTCCGTGTCGACCCATGCCGGTCTGCGCCTGTGGGA
GACGACGCTGGACCGAAAGCGGCTGCCGTGGCTCGGCGAGCACCGGGCGC
AGGGGGAGGTCGTGTTTCCTGGCGCCGGGTACCTGGAGATGGCGCTGTCG
TCGGGGGCCGAGATCTTGGGCGATGGACCGATCCAGGTCACGGATGTGGT
GCTCATCGAGACGCTGACCTTCGCGGGCGATACGGCGGTACCGGTCCAGG
TGGTGACGACCGAGGAGCGACCGGGACGGCTGCGGTTCCAGGTAGCGAGT
CGGGAGCCGGGGGAACGTCGCGCGCCCTTCCGGATCCACGCCCGCGGCGT
GCTGCGCCGGATCGGGCGCGTCGAGACCCCGGCGAGGTCGAACCTCGCCG
CCCTGCGCGCCCGGCTTCATGCCGCCGTGCCCGCTGCGGCTATCTATGGT
GCGCTCGCCGAGATGGGGCTTCAATACGGCCCGGCGTTGCGGGGGCTCGC
CGAGCTGTGGCGGGGTGAGGGCGAGGCGCTGGGCAGGGTGAGACTGCCTG
AGGCCGCCGGCTCCGCGACAGCCTACCAGCTGCATCCGGTGCTGCTGGAC
GCGTGCGTCCAAATGATTGTTGGCGCGTTCGCCGATCGCGATGAGGCGAC

GCCGTGGGCGCCGGTGGAGGTGGGCTCGGTGCGGCTGTTCCAGCGGTCTC
CTGGGGAGCTATGGTGCCATGCGCGCGTCGTGAGCGATGGTCAACAGGCC
TCCAGCCGGTGGAGCGCCGACTTTGAGTTGATGGACGGTACGGGCGCGGT
GGTCGCCGAGATCTCCCGGCTGGTGGTGGAGCGGCTTGCGAGCGGTGTAC
GCCGGCGCGACGCAGACGACTGGTTCCTGGAGCTGGATTGGGAGCCCGCG
GCGCTCGGTGGGCCCAAGATCACAGCCGGCCGGTGGCTGCTGCTCGGCGA
GGGTGGTGGGCTCGGGCGCTCGTTGTGCTCGGCGCTGAAGGCCGCCGGCC
ATGTCGTCGTCCACGCCGCGGGGGACGACACGAGCACTGCAGGAATGCGC
GCGCTCCTGGCCAACGCGTTCGACGGCCAGGCCCCGACGGCCGTGGTGCA
CCTCAGCAGCCTCGACGGGGGCGGCCAGCTCGGCCCGGGGCTCGGGGCGC
AGGGCGCGCTCGACGCGCCCCGGAGCCCAGATGTCGATGCCGATGCCCTC
GAATCGGCGCTGATGCGTGGTTGCGACAGCGTGCTCTCCCTGGTGCAAGC
GCTGGTCGGCATGGACCTCCGAAACGCGCCGCGGCTGTGGCTCTTGACCC
GCGGGGCTCAGGCGGCCGCCGCCGGCGATGTCTCCGTGGTGCAAGCGCCG
CTGTTGGGGCTGGGCCGCACCATCGCCTTGGAGCACGCCGAGCTGCGCTG
TATCAGCGTCGACCTCGATCCAGCCGAGCCTGAAGGGGAAGCCGATGCTT
TGCTGGCCGAGCTACTTGCAGATGATGCCGAGGAGGAGGTCGCGCTGCGC
GGTGGCGACCGGCTCGTTGCGCGGCTCGTCCACCGGCTGCCCGACGCTCA
GCGCCGGGAGAAGGTCGAGCCCGCCGGTGACAGGCCGTTCCGGCTAGAGA
TCGATGAACCCGGCGCGCTGGACCAACTGGTGCTCCGAGCCACGGGGCGG
CGCGCTCCTGGTCCGGGCGAGGTCGAGATCTCCGTCGAAGCGGCGGGGCT
CGACTCCATCGACATCCAGCTGGCGTTGGGCGTTGCTCCCAATGATCTGC
CTGGAGAAGAAATCGAGCCGTTGGTGCTCGGAAGCGAGTGCGCCGGGCGC
ATCGTCGCTGTGGGCGAGGGCGTGAACGGCCTTGTGGTGGGCCAGCCGGT
GATCGCCCTTGCGGCGGGAGTATTTGCTACCCATGTCACCACGTCGGCCA
CGCTGGTGTTGCCTCGGCCTCTGGGGCTCTCGGCGACCGAGGCGGCCGCG
ATGCCCCTCGCGTATTTGACGGCCTGGTACGCCCTCGACAAGGTCGCCCA
CCTGCAGGCGGGGGAGCGGGTGCTGATCCATGCGGAGGCCGGTGGTGTCG
GTCTTTGCGCGGTGCGATGGGCGCAGCGCGTGGGCGCCGAGGTGTATGCG
ACCGCCGACACGCCCGAGAACCGTGCCTACCTGGAGTCGCTGGGCGTGCG

GTACGTGAGCGATTCCCGCTCGGGCCGGTTCGTCACAGACGTGCATGCAT
GGACGGACGGCGAGGGTGTGGACGTCGTGCTCGACTCGCTTTCGGGCGAG
CGCATCGACAAGAGCCTCATGGTCCTGCGCGCCTGTGGTCGCCTTGTGAA
GCTGGGCAGGCGCGACGACTGCGCCGACACGCAGCCTGGGCTGCCGCCGC
TCCTACGGAATTTTTCCTTCTCGCAGGTGGACTTGCGGGGAATGATGCTC
GATCAACCGGCGAGGATCCGTGCGCTCCTCGACGAGCTGTTCGGGTTGGT
CGCAGCCGGTGCCATCAGCCCACTGGGGTCGGGGTTGCGCGTTGGCGGAT
CCCTCACGCCACCGCCGGTCGAGACCTTCCCGATCTCTCGCGCAGCCGAG
GCATTCCGGAGGATGGCGCAAGGACAGCATCTCGGGAAGCTCGTGCTCAC
GCTGGACGACCCGGAGGTGCGGATCCGCGCTCCGGCCGAATCCAGCGTCG
CCGTCCGCGCGGACGGCACCTACCTTGTGACCGGCGGTCTGGGTGGCCTC
GGTCTGCGCGTGGCCGGATGGCTGGCCGAGCGGGGCGCGGGGCAACTGGT
GCTGGTGGGCCGCTCCGGTGCGGCGAGCGCAGAGCAGCGAGCCGCCGTGG
CGGCGCTGGAGGCCCACGGCGCGCGCGTCACGGTGGCGAAAGCGGACGTC
GCCGATCGGTCACAGATCGAGCGGGTCCTCCGCGAGGTTACCGCGTCGGG
GATGCCGCTGCGGGGTGTCGTGCATGCGGCAGGTCTCGTGGATGACGGGC
TGCTGATGCAGCAGACTCCGGCGCGGTTCCGCACGGTGATGGGACCTAAG
GTCCAGGGGGCCTTGCACTTGCACACGCTGACACGCGAAGCGCCTCTTTC
CTTCTTCGTGCTGTACGCTTCTGCAGCTGGGCTTTTCGGCTCGCCAGGCC
AGGGCAACTATGCCGCAGCCAACGCGTTCCTCGACGCCCTTTCGCATCAC
CGAAGGGCGCAGGGCCTGCCGGCGCTGAGCATCGACTGGGGCATGTTCAC
GGAGGTGGGGATGGCCGTTGCGCAAGAAAACCGTGGCGCGCGGCAGATCT
CTCGCGGGATGCGGGGCATCACCCCCGATGAGGGTCTGTCAGCTCTGGCG
CGCTTGCTCGAGGGTGATCGCGTGCAGACGGGGGTGATACCGATCACTCC
GCGGCAGTGGGTGGAGTTCTACCCGGCAACAGCGGCCTCACGGAGGTTGT
CGCGGCTGGTGACCACGCAGCGCGCGGTCGCTGATCGGACCGCCGGGGAT
CGGGACCTGCTCGAACAGCTTGCGTCGGCTGAGCCGAGCGCGCGGGCGGG
GCTGCTGCAGGACGTCGTGCGCGTGCAGGTCTCGCATGTGCTGCGTCTCC
CTGAAGACAAGATCGAGGTGGATGCCCCGCTCTCGAGCATGGGCATGGAC
TCGCTGATGAGCCTGGAGCTGCGCAACCGCATCGAGGCTGCGCTGGGCGT

CGCCGCGCCTGCAGCCTTGGGGTGGACGTACCCAACGGTAGCAGCGATAA
CGCGCTGGCTGCTCGACGACGCCCTCGTCGTCCGGCTTGGCGGCGGTCG
GACACGGACGAATCGACGGCGAGCGCCGGTTCGTTCGTCCACGTCCTCCG
CTTTCGTCCTGTCGTCAAGCCGCGGGCTCGTCTCTTCTGTTTTCACGGTT
CTGGCGGCTCGCCCGAGGGCTTCCGTTCCTGGTCGGAGAAGTCTGAGTGG
AGCGATCTGGAAATCGTGGCCATGTGGCACGATCGCAGCCTCGCCTCCGA
GGACGCGCCTGGTAAGAAGTACGTCCAAGAGGCGGCCTCGCTGATTCAGC
ACTATGCAGACGCACCGTTTGCGTTAGTAGGGTTCAGCCTGGGTGTCCGG
TTCGTCATGGGGACAGCCGTGGAGCTCGCCAGTCGTTCCGGCGCACCGGC
TCCGCTGGCCGTCTTCACGTTGGGCGGCAGCTTGATCTCTTCTTCAGAGA
TCACCCCGGAGATGGAGACCGATATAATAGCCAAGCTCTTCTTCCGAAAT
GCCGCGGGTTTCGTGCGATCCACCCAACAAGTCCAGGCCGATGCTCGCGC
AGACAAGGTCATCACAGACACCATGGTGGCTCCGGCCCCCGGGGACTCGA
AGGAGCCGCCCGTGAAGATCGCGGTCCCTATCGTCGCCATCGCCGGCTCG
GACGATGTGATCGTGCCTCCGAGCGACGTTCAGGATCTACAATCTCGCAC
CACGGAGCGCTTCTATATGCATCTCCTTCCCGGAGATCACGAATTTCTCG
TCGATCGAGGGCGCGAGATCATGCACATCGTCGACTCGCATCTCAATCCG
CTGCTCGCCGCGAGGACGACGTCGTCAGGCCCCGCGTTCGAGGCAAAATG
ATGGCAGCCTCCCTCGGGCGCGCGAGATGGTTGGGAGCAGCGTGGGCGCT
GGCGGCCGGCGGCAGGCCGCGGAGGCGCATGAGCCTTCCTGGACGTTTGC
AGTATAGGAGATTTTATGACACAGGAGCAAGCGAATCAGAGTGAGACGAA
GCCTGCTTTCGACTTCAAGCCGTTCGCGCCTGGGTACGCGGAGGACCCGT
TCCCCGCGATCGAGCGCCTGAGAGAGGCAACCCCCATCTTCTACTGGGAT
GAAGGCCGCTCCTGGGTCCTCACCCGATACCACGACGTGTCGGCGGTGTT
CCGCGACGAACGCTTCGCGGTCAGTCGAGAAGAGTGGGAATCGAGCGCGG
AGTACTCGTCGGCCATTCCCGAGCTCAGCGATATGAAGAAGTACGGATTG
TTCGGGCTGCCGCCGGAGGATCACGCTCGGGTCCGCAAGCTCGTCAACCC
GTCGTTTACGTCACGCGCCATCGACCTGCTGCGCGCCGAAATACAGCGCA
CCGTCGACCAGCTGCTCGATGCTCGCTCCGGACAAGAGGAGTTCGACGTT
GTGCGGGATTACGCGGAGGGAATCCCGATGCGCGCGATCAGCGCTCTGTT

GAAGGTTCCGGCCGAGTGTGACGAGAAGTTCCGTCGCTTCGGCTCGGCGA
CTGCGCGCGCGCTCGGCGTGGGTTTGGTGCCCCAGGTCGATGAGGAGACC
AAGACCCTGGTCGCGTCCGTCACCGAGGGGCTCGCGCTGCTCCATGACGT
CCTCGATGAGCGGCGCAGGAACCCGCTCGAAAATGACGTCTTGACGATGC
TGCTTCAGGCCGAGGCCGACGGCAGCAGGCTGAGCACGAAGGAGCTGGTC
GCGCTCGTGGGTGCGATTATCGCTGCTGGCACCGATACCACGATCTACCT
TATCGCGTTCGCTGTGCTCAACCTGCTGCGGTCGCCCGAGGCGCTCGAGC
TGGTGAAGGCCGAGCCCGGGCTCATGAGGAACGCGCTCGATGAGGTGCTC
CGCTTCGACAATATCCTCAGAATAGGAACTGTGCGTTTCGCCAGGCAGGA
CCTGGAGTACTGCGGGGCATCGATCAAGAAAGGGGAGATGGTCTTTCTCC
TGATCCCGAGCGCCCTGAGAGATGGGACTGTATTCTCCAGGCCAGACGTG
TTTGATGTGCGACGGGACACGGGCGCGAGCCTCGCGTACGGTAGAGGCCC
CCATGTCTGCCCCGGGGTGTCCCTTGCTCGCCTCGAGGCGGAGATCGCCG
TGGGCACCATCTTCCGTAGGTTCCCCGAGATGAAGCTGAAAGAAACTCCC
GTGTTTGGATACCACCCCGCGTTCCGGAACATCGAATCACTCAACGTCAT
CTTGAAGCCCTCCAAAGCTGGATAGCTCGCGGGGGTATCGCTTCCCGAAC
CTCATTCCCTCATGATACAGCTCGCGCGCGGGTGCTGTCTGCCGCGGGTG
CGATTCGATCCAGCGGACAAGCCCATTGTCAGCGCGCGAAGATCGAATCC
ACGGCCCGGAGAAGAGCCCGTCCGGGTGACGTCGGAAGAAGTGCCGGGCG
CCGCCCTGGGAGCGCAAAGCTCGCTCGTTCGCGCTCAGCACGCCGCTCGT
CATGTCCGGCCCTGCACCCGCGCCGAGGAGCCGCCCGCCCTGATGCACGG
CCTCACCGAGCGGCAGGTTCTGCTCTCGCTCGTCGCCCTCGCGCTCGTCC
TCCTGACCGCGCGCGCCTTCGGCGAGCTCGCGCGGCGGCTGCGCCAGCCC
GAGGTGCTCGGCGAGCTCTTCGGCGGCGTGGTGCTGGGCCCGTCCGTCGT
CGGCGCGCTCGCTCCTGGGTTCCATCGAGTCCTCTTCCAGGATCCGGCGG
TCGGGGTCGTGCTCTCCGGCATCTCCTGGATAGGCGCGCTCGTCCTGCTG
CTCATGGCGGGTATCGAGGTCGATGTGAGCATCCTGCGCAAGGAGGCGCG
CCCCGGGGCGCTCTCGGCGCTCGGCGCGATCGCGCCCCCGCTGCGCACGC
CGGGGCCGCTGGTGCAGCGCATGCAGGGCGCGTTCACGTGGGATCTCGAC
GTCTCGCCGCGACGCTCTGCGCAAGCCTGAGCCTCGGCGCCTGCTCGTAC

ACCTCGCCGGTGCTCGCTCCGCCCGCGGACATCCGGCCGCCCGCCGCGGC
CCAGCTCGAGCCGGACTCGCCGGATGACGAGGCCGACGAGGCCGACGAGG
CGCTCCGCCCGTTCCGCGACGCGATCGCCGCGTACTCGGAGGCCGTTCGG
TGGGCGGAGGCGGCGCAGCGGCCGCGGCTGGAGAGCCTCGTGCGGCTCGC
GATCGTGCGGCTGGGCAAGGCGCTCGACAAGGTCCCTTTCGCGCACACGA
CGGCCGGCGTCTCCCAGATCGCCGGCAGACTCCAGAACGATGCGGTCTGG
TTCGATGTCGCCGCCCGGTACGCGAGCTTCCGCGCGGCGACGGAGCACGC
GCTCCGCGACGCGGCGTCGGCCATGGAGGCGCTCGCGGCCGGCCCGTACC
GCGGATCGAGCCGCGTGTCCGCTGCCGTAGGGGAGTTTCGGGGGGAGGCG
GCGCGCCTTCACCCCGCGGACCGTGTACCCGCGTCCGACCAGCAGATCCT
GACCGCGCTGCGCGCAGCCGAGCGGGCGCTCATCGCGCTCTACACTGCGT
TCGCCCGTGAGGAGTGAGCCTCTCTCGGGCGCAGCCGAGCGGCGGCGTGC
CGGTGGTTCCCTCTTCGCAACCATGACCGGAGCCGCGCTCGGTCCGCGCA
GCGGCTAGCGCGCGTCGCGGCAGAGATCGCTGGAGCGACAGGCGACGACC
CGCCCGAGGGTGTCGAACGGATTGCCGCAGCCCTCATTGCGGATCCCCTC
CAGACACTCGTTCAGCTGCTTGGCGTCGATGCCGCCTGGGCACTCGCCGA
AGGTCAGCTCGTCGCGCCACTCGGATCGGATCTTGTTCGAGCACGCGTCC
TTGCTCGAATACTCCCGGTCTTGTCCGATGTTGTTGCACCGCGCCTCGCG
GTCGCACCGCGCCGCCACGATGCTATCGACGGCGCTGCCGACTGGCACCG
GCGCCTCGCCCTGCGCGCCACCCGGGGTTTGCGCCTCCCCGCCTGACCGC
TTTTCGCCGCCGCACGCCGCGAGCAGGCTCATTCCCGACACCGAGATCAG
GCCCACGACCAGCTTCCCAGCAATCTTTTGCATGGCTTCCCCTCCCTCAC
GACACGTCACATCAGAGACTCTCCGCTCGGCTCGTCGGTTCGACAGCCGG
CGACGGCCACGAGCAGAACCGTCCCCGACCAGAACAGCCGCATGCGGGTT
TCTCGCAACATGCCCCGACATCCTTGCGACTAGCGTGCCTCCGCTCGTGC
CGAGATCGGCTGTCCTGTGCGACGGCAATATCCTGCGATCGGCCGGGCAG
GAGGTACCGACACGGGCGCCGGGCGGGAGGTGCCGCCACGGGCTCGAAAT
GTGCTGCGGCAGGCGCCTCCATGCCCGCAGCCGGGAACGCGGCGCCCGGC
CAGCCTCGGGGTGACGCCGCAAACGGAGATGCTCCCGGAGAGGCGCCGG
GCACAGCCGAGCGCCGTCACCACCGTGCGCACTCGTGAGCTCCAGCTCCT

CGGCATAGAAGAGACCGTCACTCCCGGTCCGTGTAGGCGATCGTGCTGAT
CAGCGCGTTCTCCGCCTGACGCGAGTCGAGCCGGGTATGCTGCACGACAA
TGGGAACGTCCGATTCGATCACGCTGGCATAGTCCGTATCGCGCGGGATC
GGCTCGGGTTCGGTCAGATCGTTGAACCGGACGTGCCGGGTGCGCCTCGC
TGGGACGGTCACCCGGTACGGCCCGGCGGGGTCGCGGTCGCTGAAGTAGA
CGGTGATGGCGACCTGCGCGTCCCGGTCCGACGCATTCAACAGGCAGGCC
GTCTCATGGCTCGTCATCTGCGGCTCGGGTCCGTTGCTCCGGCCTGGGAT
GTAGCCCTCTGCGATTGCCCAGCGCGTCCGCCCGATCGGCTTCTCCATAT
GTCCTCCCTGCTGGCTCCTCTTTGGCTGCCTCCCTCTGCTGTCCAGGAGC
GACGGCCTCTTCTCCCGACGCGCTCGGGGATCCATGGCTGAGGATCCTCG
CCGAGCGCTCCTTGCCGACCGGCGCGCCGAGCGCCGACGGGCTTTGAAAG
CACGCGACCGGACACGTGATGCCGGCGCGACGAGGCCGCCCCGCGTCTGA
TCCCGATCGTGACATCGCGACGTCCGCCGGCGCCTCTGCAGGCCGGCCTG
AGCGTTGCGCGGTCATGGTCGTCCTCGCGTCACCGCCACCCGCCGATTCA
CATCCCACCGCGGCACGACGCTTGCTCAAACCGCGGCGAGACGGCCGGGC
GGCTGTGGTACCGGCCAGCCCGGACGCGAGGCCCGAGAGGGACAGTGGGT
CCGCCGTGAAGCAGTGAGGCGATCGAGGTGGCAGATGAAACACGTTGACA
CGGGCCGACGAGTCGGCCGCCGGATAGGGCTCACGCTCGGTCTCCTCGCG
AGCATGGCGCTCGCCGGCTGTGGCGGCCCGAGCGAGAAAATCGTGCAGGG
CACGCGGCTCGCGCCCGGCGCCGATGCGCACGTCGCCGCCGACGTCGACC
CCGACGCCGCGACCACGCGGCTGGCGGTGGACGTCGTTCACCTCTCGCCG
CCCGAGCGCATCGAGGCCGGCAGCGAGCGGTTCGTCGTCTGGCAGCGTCC
GAGCTCCGAGTCCCCGTGGCAACGGGTCGGAGTGCTCGACTACAACGCTG
CCAGCCGAAGAGGCAAGCTGGCCGAGACGACCGTGCCGCATGCCAACTTC
GAGCTGCTCATCACCGTCGAGAAGCAGAGCAGCCCTCAGTCTCCATCTTC
TGCCGCCGTCATCGGGCCGACGTCCGTCGGGTAACATCGCGCTATCAGCA
GCGCTGAGCCCGCCAGCAGGCCCCAGAGCCCTGCCTCGATCGCCTTCTCC
ATCATATCATCCCTGCGTACTCCTCCAGCGACGGCCGCGTCGAAGCAACC
GCCGTGCCGGCGCGGCTCTACGTGCGCGACAGGAGAGCGTCCTGGCGCGG
CCTGCGCATCGCTGGAAGGATCGGCGGAGCATGGAGAAAGAATCGAGGAT

CGCGATCTACGGCGCCATCGCAGCCAACGTGGCGATCGCGGCGGTCAAGT
TCATCGCCGCCGCCGTGACCGGCAGCTCGGCGATGCTCTCCGAGGGCGTG
CACTCCCTCGTCGATACTGCAGACGGGCTCCTCCTCCTGCTCGGCAAGCA
CCGGAGCGCACGCCCGCCCGACGCCGAGCATCCGTTCGGCCACGGCAAGG
AGCTCTATTTCTGGACGCTGATCGTCGCCATCATGATCTTCGCCGCGGGC
GGCGGCGTCTCGATCTACGAAGGGATCTTGCACCTCTTGCACCCGCGCCA
GATCGAGGATCCGACGTGGAACTACGTCGTCCTCGGCGCAGCGGCCGTCT
TCGAGGGGACGTCGCTCATCATCTCGATCCACGAGTTCAAGAAGAAGGAC
GGACAGGGCTACCTCGCGGCGATGCGGTCCAGCAAGGACCCGACGACGTT
CACGATCGTCCTGGAGGACTCCGCGGCGCTCGCCGGGCTCACCATCGCCT
TCCTCGGCGTCTGGCTCGGGCACCGCCTGGGAAACCCCTACCTCGACGGC
GCGGCGTCGATCGGCATCGGCCTCGTGCTCGCCGCGGTCGCGGTCTTCCT
CGCCAGCCAGAGCCGTGGGCTCCTCGTGGGGAGAGCGCGGACAGGGAGC
TCCTCGCCGCGATCCGCGCGCTCGCCAGCGCAGATCCTGGCGTGTCGGCG
GTGGGGCGGCCCCTGACGATGCACTTCGGTCCGCACGAAGTCCTGGTCGT
GCTGCGCATCGAGTTCGACGCCGCGCTCACGGCGTCCGGGGTCGCGGAGG
CGAGGGAGCGCATCGAGACCCGGATACGGAGCGAGCGACCCGACGTGAAG
CACATCTACGTCGAGGCCAGGTCGCTCCACCAGCGCGCGAGGGCGTGACG
CGCCGTGGAGAGACCGCGCGCGGCCTCCGCCATCCTCCGCGGCGCCCGGG
CTCAGGTGGCCCTCGCAGCAGGGCGCGCCTGGCGGGCAAACCGTGCAGAC
GTCGTCCTTCGACGCGAGGTACGCTGGTTGCAAGTCGTCACGCCGTATCG
CGAGGTCCGGCAGCGCCGGAGCCCGGGCGGGCCGGGCGCACGAAGGCGCG
GCGAGCGCAGGCTTCGAGGGGGGCGACGTCATGAGGAAGGCCAGGGCGCA
TGGGGCGATGCTCGGCGGGCGAGATGACGGCTGGCGTCGCGGCCTCCCCG
GCGCCGGCGCGCTTCGCGCCGCGCTCCAGCGCGGTCGCTCGCGCGATCTC
GCCCGGCGCCGGCTCATCGCCTCCGTGTCCCTCGCCGGCGGCGCCAGCAT
GGCGGTCGTCTCGCTGTTCCAGCTCGGGATCATCGAGCGCCTGCCCGATC
CTCCGCTTCCAGGGTTCGATTCGGCCAAGGTGACGAGCTCCGATATCGCG
TTCGGGCTCACGATGCCGGACGCGCCGCTCGCGCTCACCAGCTTCGCGTC
CAACCTCGCGCTGGCTGGCTGGGGAGGCGCCGAGCGCGCCAGGAACACCC

CCTGGATCCCCGTCGCCGTGGCGGCCAAGGCGGCCGTCGAGGCGGCCGTG
TCCGGATGGCTCCTCGTCCAGATGCGACGGCGGGAGAGGGCCTGGTGCGC
GTACTGCCTGGTCGCCATGGCGGCCAACATGGCCGTGTTCGCGCTCTCGC
TCCCGGAAGGGTGGGCGGCGCTGGGGAAGGCGCGAGCGCGCTCGTGACAG
GACGGGCGCGGGCAGCCCCGGCCATCGGAGGCCGGCGTGCACCCGCTCCG
TCACGCCCCAGCCCGCGCCGCGTGATCTCCCGCGGACAGGGCGCGTACCG
TGGACCCCGCACGCGCCGCGTCGACGGACATCCCCGGCGACCCGCGCGGC
GCGACCCGCGCAACTCCGGCCCGCCGCCGGGCATCGACATCTCCCGTGAG
CAAGGGCACTCCGCTCCTGCCCGCGTCCGCGAACGATGGCTGCGCTGTTT
CCACCCTGGAGCAACTCCGTTTACCGCGTGGCGCTCGTCGGGCTCGTCGC
CTCGGCGGGCGGCGCCATCCTCGCGCTCATGATCTACGTCCGCACGCCGT
GGAAGCGATACCAGTTCGAGCCCGTCGATCAGCCGGTGCAGTTCGATCAC
CGCCATCACGTGCAGGACGACGCCATCGATTGCGTCTACTGCCACACCAC
GGTGACCCGCTCGCCCACGGCGGGGATGCCGCCGACGGCCACGTGCATGG
GGTGCCACAGCCAGATCTGGAATCAGAGCGTCATGCTCGAGCCCGTGCGG
CGGAGCTGGTTCTCCGGCCACGCCGATCCCGTGGAACCGGGTGAAACTCC
GTGCCCGACTTCGTCTATTTCAACCACGCGATCCACGTGAACAAGGGCGT
GGGCTGGCGTGAAGCTGCCACGGGCGCGTGGACGAGATGGCGGCCGTCTA
CAAGGTGGCGCCGATGACGATGGGCTGGTGCCTGGAGTGCCATCGCCTGC
CGGAGCCGCACCTCCGCCCGCTCTCCGCGATCACCGACATGCGCTGGGAC
CCGGGGGAGCGGAGGGATGAGCTCGGGGCGCAGCTCGCGAAGGAATACGG
GGTCCGGCGGCTCACGCACTGCACAGCGTGCCATCGATGAACGATGAACA
GGGGATCTCCTTGAAAGACGCAGATGAGATGAAGGAATGGTGGCTAGAAG
CGCTCGGGCCGGCGGGAGAGCGCGCGTCCTACAGGCTGCTGGCGCCGCTC
ATCGAGAGCCCGGAGCTCCGCGCGCTCGCCGCGGGCGAACCGCCCCGGGG
CGTGGACGAGCCGGCGGGCGTCAGCCGCCGCGCGCTGCTCAAGCTGCTCG
GCGCGAGCATGGCGCTCGCCGGCGTCGCGGGCTGCACCCCGCATGAGCCC
GAGAAGATCCTGCCGTACAACGAGACCCCGCCCGGCGTCGTGCCGGGTCT
CTCCCAATCCTACGCGACGAGCATGGTGCTCGACGGGTATGCCATGGGCC
TCCTCGCCAAGAGCTACGCGGGGCGGCCCATCAAGATCGAGGGCAACCCC

GCGCACCCGGCGAGCCTCGGCGCGACCGGCGTCCACGAGCAGGCCTCGAT
CCTCTCGCTGTACGACCCGTACCGCGCGCGCGCCGACGCGCGGCGGCC
AGGTCGCGTCGTGGGAGGCGCTCTCCGCGCGCTTCGGCGGCGACCGCGAG
GACGGCGGCGCTGGCCTCCGCTTCGTCCTCCAGCCCACGAGCTCGCCCCT
CATCGCCGCGCTGATCGAGCGCGTCCGGCGCAGGTTCCCCGGCGCGCGGT
TCACCTTCTGGTCGCCGGTCCACGCCGAGCACGCGCTCGAAGGCGCGCGG
GCGGCGCTCGGCCTCAGGCTCTTGCCTCAGCTCGACGTCGACCAGGCCGA
GGTGATCCTCTCGCTGGACGCGGACTTCCTCGCGGACATGCCGTTCAGCG
TGCGCTATGCGCGCGACTTCGCCGCGCGCCGCCGCCCCGCGAGCCCGGCG
GCGGCCATGAGCCGCCTCTACGTCGCGGAGGCGATGTTCACGCCCACGGG
GACGCTCGCCGACCACCGGCTCCGCGTGCGGCCCGCCGAGGTCGCGCGCG
TCGCGGCCGGCGTCGCGGTGGGAACTCGTGCACGAGTCTTTGTCTTGCGC
CCTGTCCGGGAATAACGGACACCTTATCGCGGGTCGCTCTTTGTGCGCGG
CTTCTGTACCTCTCAGGACAGGTAGAAGAGGGACTCAGGGGCCCTTATGT
TAACTGGGGATGCCTTCGGGACGGCCGCAAATATATCCTATCACCTCACT
GGGTGTGGGGGAGCACCGCGAGGATGTACAACCTCTGTAACTCTATGTGA
GATAATGTGTGCAGTGATCTGAGACTTATTTGTGTGACCGAGACGTCTCT
CTTATTGGTACGCATAGTATAATATAACACGTCTCATACATACTCCCGAC
ATATCCGCGGTATGCGCGCACATAGAATAGGTGATGATAAATCCCTAGTG
TGTGGAACTAGAAGATGCGGGAGTTACCTGATATTTACGGAAAAAGTATT
ATCTCAACTACCTCTCTGTTGAGACTATCACTTCGGTGTCGTTGTGCTGC
TGGT,

or its complementary strand,

(b) DNA-sequences which hybridise under stringent conditions to regions of DNA-sequences according to (a) encoding proteins or to fragments of said DNA-sequences,

(c) DNA-sequences which hybridise to the DNA-sequences according to (a) and (b) because of a degeneration of the genetic code,

(d) allele variations and mutants resulting by substitution, insertion or deletion of nucleotides or inversion of nucleotide segments of DNA-sequences according to (a) to (c), wherein the variations and mutants offer isofunctional expression products.

11. Peptide encoded by a DNA sequence according to claim 10 selected from the group consisting of

Seq ID No 83

```
>Contig56_003 2890 amino acids MW=307428 D pI=5.76 numambig=13
IRPRAAAVPMRSTVTGGVIAGPELGASYWADNLRQPVRFAAAAQALLEGGPALFIEMSPH
PILVPPLDEIQTAAEQGGAAVGSLRRGQDERATLLEALGTLWASGYPVSWARLFPAGGRR
VPLPTYPWQHERCWIEVEPDARRLAAADPTKDWFYRTDWPEVPRAAPKSETAHGSWLLLA
DRGGVGEAVAAALSTRGLSCTVLHASADASTVAEQVSEAASRRNDWQGVLYLWGLDAVVD
AGASADEVSEATRRATAPVLGLVRFLSAAPHPPRFWVVTRGACTVGGEPEASLCQAALWG
LARVAALEHPAAWGGLVDLDPQKSPTEIEPLVAELLSPDAEDQLAFRSGRRHAARLVAAP
PEGDVAPISLSAEGSYLVTGGLGGLGLLVARWLVERGARHLVLTSRHGLPERQASGGEQP
PEARARIAAVEGLEAQGARVTVAAVDVAEADPMTALLAAIEPPLRGVVHAAGVFPVRHLA
ETDEALLESVLRPKVAGSWLLHRLLRDRPLDLFVLFSSGAAVWGGKGQGAYAAANAFLDG
LAHHRRAHSLPALSLAWGLWAEGGMVDAKAHARLSDIGVLPMATGPALSALERLVNTSAV
QRSVTRMDWARFAPVYAARGRRNLLSALVAEDERAASPPVPTANRIWRGLSVAESRSALY
ELVRGIVARVLGFSDPGALDVGRGFAEQGLDSLMALEIRNRLQRELGERLSATLAFDHPT
VERLVAHLLTDVLKLEDRSDTRHIRSVAADDDIAIVGAACRFPGGDEGLETYWRHLAEGM
VVSTEVPADRWRAADWYDPDPEVPGRTYVAKGAFLRDVRSLDAAFFAISPREAMSLDPQQ
RLLLEVSWEAIERAGQDPMALRESATGVFVGMIGSEHAERVQGLDDDAALLYGTTGNLLS
VAAGRLSFFLGLHGPTMTVDTACSSSLVALHLACQSLRLGECDQALAGGSSVLLSPRSFV
```

AASRMRLLSPDGRCKTFSAAADGFARAEGCAVVVLKRLRDAQRDRDPILAVVRSTAINHD
GPSSGLTVPSGPAQQALLRQALAQAGVAPAEVDFVECHGTGTALGDPIEVQALGAVYGRG
RPAERPLWLGAVKANLGHLEAAAGLAGVLKVLLALEHEQIPAQPELDELNPHIPWAELPV
AVVRRAVPWPRGARPRRAGVSAFGLSGTNAHVVLEEAPAVEPVAAAPERAAELFVLSAKS
AAALDAQAARLRDHLEKHVELGLGDVAFSLATTRSAMEHRLAVAASSREALRGALSAAAQ
GHTPPGAVRGRASGGSAPKVVFVFPGQGSQWVGMGRKLMAEEPVFRAALEGCDRAIEAEA
GWSLLGELSADEAASQLGRIDVVQPVLFAMEVALSALWRSWGVEPEAVVGHSMGEVAAAH
VAGALSLEDAVAIICRRSRLLRRISGQGEMALVELSLEEAEAALRGHEGRLSVAVSNSPR
STVLAGEPAALSEVLAALTAKGVFWRQVKVDVASHSPQVDPLREELIAALGAIRPRAAAV
PMRSTVTGGVIAGPELGASYWADNLRQPVRFAAAAQALLEGGPALFIEMSPHPILVPPLD
EIQTAAEQGGAAVGSLRRGQDERATLLEALGTLWASGYPVSWARLFPAGGRRVPLPTYPW
QHERYWIEDSVHGSKPSLRLRQLRNGATDHPLLGAPLLVSARPGAHLWEQALSDERLSYL
SEHRVHGEAVLPSAAYVEMALAAGVDLYGTATLVLEQLALERALAVPSEGGRIVQVALSE
EGPGRASFQVSSREEAGRSWVRHATGHVCSGQSSAVGALKEAPWEIQRRCPSVLSSEALY
PLLNEHALDYGPCFQGVEQVWLGTGEVLGRVRLPGDMASSSGAYRIHPALLDACFQVLTA
LLTTPESIEIRRRLTDLHEPDLPRSRAPVNQAVSDTWLWDAALDGGRRQSASVPVDLVLG
SFHAKWEVMERLAQAYIIGTLRIWNVFCAAGERHTIDELLVRLQISVVYRKVIKRWMEHL
VAIGILVGDGEHFVSSQPLPEPDLAAVLEEAGRVFADLPVLFEWCKFAGERLADVLTGKT
LALEILFPGGSFDMAERIYRDSPIARYSNGIVRGVVESAARVVAPSGMFSILEIGAGTGA
TTAAVLPVLLPDRTEYHFTDVSPLFLARAEQRFRDYPFLKYGILDVDQEPAGQGYAHQRF
DVIVAANVIHATRDIRATAKRLLSLLAPGGLLVLVEGTGHPIWFDITTGLIEGWQKYEDD
LRIDHPLLPARTWCDVLRRVGFADAVSLPGDGSPAGILGQHVILSRAPGIAGAACDSSGE
SATESPAARAVRQEWADGSADVVHRMALERMYFHRRPGRQVWVHGRLRTGGGAFTKALAG
DLLLFEDTGQVVAEVQGLRLPQLEASAFAPRDPREEWLYALEWQRKDPIPEAPAAASSSS
AGAWLVLMDQGGTGAALVSLLEGRGEACVRVIAGTAYACLAPGLYQVDPAQPDGFHTLLR
DAFGEDRICRAVVHMWSLDATAAGERATAESLQADQLLGSLSALSLVQALVRRRWRNMPR
LWLLTRAVHAVGAEDAAASVAQAPVWGLGRTLALEHPELRCTLVDVNPAPSPEDAAALAV
ELGASDREDQVALRSDGRYVARLVRSSFSGKPATDCGIRADGSYVITDGMGRVGLSVAQW
MVMQGARHVVLVDRGGASEASRDALRSMAEAGAEVQIVEADVARRDDVARLLSKIEPSMP
PLRGIVYVDGTFQGDSSMLELDARRFKEWMYPKVLGAWNLHALTRDRSLDFFVLYSSGTS

LLGLPGQGSRAAGDAFLDAIAHHRCKVGLTAMSINWGLLSEASSPATPNDGGARLEYRGM
EGLTLEQGAAALGRLLARPRAQVGVMRLNLRQWLEXXXXXXXXXXXXXWYNLLIIIQYTK
VPFQGPFRML*

Seq ID No 84

>Contig56_027 700 amino acids MW=80569 D pI=7.02 numambig=0
MNMELNITSKSNPFGDTTAENDKKMLSNAFIETADFRTLIETDDRTIVVGRRGTGKSALF
IQLNEHWKKDKKILILSFSPDDSQIIGFRSMLKPFTGSFNLARAATRLLWRYAMLMEIAS
YISSHYKLSSQISSETLLNEHLKKWNSAQGDILRKCRLVAKEYLDENNPEESIGDLQFNL
NISEIENNIVSLLERSDRKVVILMDKLDEAYEPDNIGIGIIAGLAYASIELNQKAKCIRP
IIFLRDNIFRSLSKEDPDYSRNIEGQVIRLHWDWAQLLMLSAKRMKVAFKLDIEKDQRVW
DRCTADDLKGRNGFKRCLQFTLYRPRDLLSLLNEAFFSAFRENRETIINTDLEYAAKSIS
MARLEDLWKEYQKIFPSIQVITSAFRSIEPELTVYTCLKKIEASFELIEENGDPKITSEI
QLLKASGILQSLYSVGFVGIRDKNTSSYSFCHDGRTPDKGFESNEKLLIHPCYWLGLNLN
RNALAPEEAEEINDEYDINIISDNSAIRNKTIGQITTHLDQIPIGNEGATEFEQWCLDAL
RIVFASHLTDIKSHPNGNAVQRRDIIGTNGGKSDFWKRVLEDYKTRQVVFDAKNFEELGP
SEYRQLQSYLTGPYGKLGFIINRDESEVLKSGKDLDWTKEMYQSHNSLIIKLPAKYISKL
LQKLRNPEKHDAIDRQMGKLLTLYETSYMAIKSTQKKRRK*

Seq ID No 85

>Contig57_001 372 amino acids MW=38411 D pI=12.39 numambig=10
MLTSXXXXXXXXXXLLAYRCATAARGAGRIRHHRQYAQRRRRTVVLYARAAGTLPDRRHG
LLVVARGHPPCLPQPARSRERSRAGGRRQHAPFVQDDDNAGAHPGAVARWPLPDIRRLGQ
RVRPWGGLRYGRAQTALRRPATRRSDLGSDPGFGHESGWPVDRVDGTQCARSGGALARGA
AERSRRRRGHRLCRDPRNGDLARRPDRGRGAACRVGAGAGRWEPLRAGRSEDKPRPPGGR
CRRGGFDQGGAGSAPRTDPAKPPFPHAQSADPDRGDRARAGDGAGAVAAGGPTALRGGER
VRPQRHQRPCRAGGGAGHGARTGDAGALSGAFGAVGEERRRAGRTGGAALSAHRRVPGAG
SRRRRVQPGIDA*

Seq ID No 86

```
>Contig57_002 2259 amino acids MW=238258 D pI=5.92 numambig=0
MSYTLGLQGPCLTVDTVCSSSLVAIHLACRSLRARESDLALAGGVNMLLSSKTMIMLGRI
QALSPDGHCRTFDASANGFVRGEGCGMVVLKRLSDAQRHGDRIWALIRGSAMNQDGRSTG
LMAPNVLAQEALLREALQSARVDAGAIGYVETHGTGTSLGDPIEVEALRAVLGPARADGS
RCVLGAVKTNLGHLEGAAGVAGLIKAALALHHELIPRNLHFHTLNPRIRIEGTALALATE
PVPWPRAGRPRFAGVSAFGLSGTNVHVVLEEAPATVLAPATPGRSAELLVLSAKSAAALD
AQAARLSAHIAAYPEQGLGDVAFSLVSTRSPMEHRLAVAATSREALRSALEVAAQGQTPA
GAARGRAASSPGKLAFLFAGQGAQVPGMGRGLWEAWPAFRETFDRCVTLFDRELHQPLCE
VMWAEPGSSRSSLLDQTAFTQPALFALEYALAALFRSWGVEPELVAGHSLGELVAACVAG
VFSLEDAVRLVVARGRLMQALPAGGAMVSIAAPEADVAAAVAPHAALVSIAAVNGPEQVV
IAGAEKFVQQIAAAFAARGARTKPLHVSHAFHSPLMDPMLEAFRRVTESVTYRRPSIALV
SNLSGKPCTDEVSAPGYWVRHAREAVRFADGVKALHAAGAGLFVEVGPKPTLLGLVPACL
PDARPVLLPASRAGRDEAASALEALGGFWVVGGSVTWSGVFPSGGRRVPLPTYPWQRERY
WIEAPVDREADGTGRARAGGHPLLGEVFSVSTHAGLRLWETTLDRKRLPWLGEHRAQGEV
VFPGAGYLEMALSSGAEILGDGPIQVTDVVLIETLTFAGDTAVPVQVVTTEERPGRLRFQ
VASREPGERRAPFRIHARGVLRRIGRVETPARSNLAALRARLHAAVPAAAIYGALAEMGL
QYGPALRGLAELWRGEGEALGRVRLPEAAGSATAYQLHPVLLDACVQMIVGAFADRDEAT
PWAPVEVGSVRLFQRSPGELWCHARVVSDGQQASSRWSADFELMDGTGAVVAEISRLVVE
RLASGVRRRDADDWFLELDWEPAALGGPKITAGRWLLLGEGGGLGRSLCSALKAAGHVVV
HAAGDDTSTAGMRALLANAFDGQAPTAVVHLSSLDGGGQLGPGLGAQGALDAPRSPDVDA
DALESALMRGCDSVLSLVQALVGMDLRNAPRLWLLTRGAQAAAAGDVSVVQAPLLGLGRT
IALEHAELRCISVDLDPAEPEGEADALLAELLADDAEEEVALRGGDRLVARLVHRLPDAQ
RREKVEPAGDRPFRLEIDEPGALDQLVLRATGRRAPGPGEVEISVEAAGLDSIDIQLALG
VAPNDLPGEEIEPLVLGSECAGRIVAVGEGVNGLVVGQPVIALAAGVFATHVTTSATLVL
PRPLGLSATEAAAMPLAYLTAWYALDKVAHLQAGERVLIHAEAGGVGLCAVRWAQRVGAE
VYATADTPENRAYLESLGVRYVSDSRSGRFVTDVHAWTDGEGVDVVLDSLSGERIDKSLM
VLRACGRLVKLGRRDDCADTQPGLPPLLRNFSFSQVDLRGMMLDQPARIRALLDELFGLV
AAGAISPLGSGLRVGGSLTPPPVETFPISRAAEAFRRMAQGQHLGKLVLTLDDPEVRIRA
PAESSVAVRADGTYLVTGGLGGLGLRVAGWLAERGAGQLVLVGRSGAASAEQRAAVAALE
```

AHGARVTVAKADVADRSQIERVLREVTASGMPLRGVVHAAGLVDDGLLMQQTPARFRTVM
GPKVQGALHLHTLTREAPLSFFVLYASAAGLFGSPGQGNYAAANAFLDALSHHRRAQGLP
ALSIDWGMFTEVGMAVAQENRGARQISRGMRGITPDEGLSALARLLEGDRVQTGVIPITP
RQWVEFYPATAASRRLSRLVTTQRAVADRTAGDRDLLEQLASAEPSARAGLLQDVVRVQV
SHVLRLPEDKIEVDAPLSSMGMDSLMSLELRNRIEAALGVAAPAALGWTYPTVAAITRWL
LDDALVVRLGGGSDTDESTASAGSFVHVLRFRPVVKPRARLFCFHGSGGSPEGFRSWSEK
SEWSDLEIVAMWHDRSLASEDAPGKKYVQEAASLIQHYADAPFALVGFSLGVRFVMGTAV
ELASRSGAPAPLAVFTLGGSLISSSEITPEMETDIIAKLFFRNAAGFVRSTQQVQADARA
DKVITDTMVAPAPGDSKEPPVKIAVPIVAIAGSDDVIVPPSDVQDLQSRTTERFYMHLLP
GDHEFLVDRGREIMHIVDSHLNPLLAARTTSSGPAFEAK*

Seq ID No 87

>Contig57_027 419 amino acids MW=46737 D pI=5.09 numambig=0
MTQEQANQSETKPAFDFKPFAPGYAEDPFPAIERLREATPIFYWDEGRSWVLTRYHDVSA
VFRDERFAVSREEWESSAEYSSAIPELSDMKKYGLFGLPPEDHARVRKLVNPSFTSRAID
LLRAEIQRTVDQLLDARSGQEEFDVVRDYAEGIPMRAISALLKVPAECDEKFRRFGSATA
RALGVGLVPQVDEETKTLVASVTEGLALLHDVLDERRRNPLENDVLTMLLQAEADGSRLS
TKELVALVGAIIAAGTDTTIYLIAFAVLNLLRSPEALELVKAEPGLMRNALDEVLRFDNI
LRIGTVRFARQDLEYCGASIKKGEMVFLLIPSALRDGTVFSRPDVFDVRRDTGASLAYGR
GPHVCPGVSLARLEAEIAVGTIFRRFPEMKLKETPVFGYHPAFRNIESLNVILKPSKAG*

Seq ID No 88

>Contig57_043 492 amino acids MW=52617 D pI=11.54 numambig=0
MAARARKSCRARGSRPAPMRTSPPTSTPTPRPRGWRWTSFTSRRPSASRPAASGSSSGSV
RAPSPRGNGSECSTTTLPAEEASWPRRPCRMPTSSCSSPSRSRAALSLHLLPPSSGRRPS
GNIALSAALSPPAGPRALPRSPSPSYHPCVLLQRRPRRSNRRAGAALRARQESVLARPAH
RWKDRRSMEKESRIAIYGAIAANVAIAAVKFIAAAVTGSSAMLSEGVHSLVDTADGLLLL
LGKHRSARPPDAEHPFGHGKELYFWTLIVAIMIFAAGGGVSIYEGILHLLHPRQIEDPTW
NYVVLGAAAVFEGTSLIISIHEFKKKDGQGYLAAMRSSKDPTTFTIVLEDSAALAGLTIA
FLGVWLGHRLGNPYLDGAASIGIGLVLAAVAVFLASQSRGLLVGESADRELLAAIRALAS

ADPGVSAVGRPLTMHFGPHEVLVVLRIEFDAALTASGVAEARERIETRIRSERPDVKHIY
VEARSLHQRARA*

12. DNA sequence according to any of claims 1 to 5 wherein the DNA is selected from the group consisting of

(a) the following DNA sequences:

Seq ID No 89 (>Contig10)

GGTAGTGAAATATGCTGTATTCAACAGAAAGCTTGATGAATTGATCTAGA
AAGTAGAGCGAGAGAATCAAGTAAGATAGTAGGATGCATTATAAATATAG
AATATATACTGCATACGATGACAGCATGCGCACGAATAGAATGCATAAGA
GGCAAGCCAATAACCAAAAGTGGAGCCAGAGGAGATAGTCTCGCCAGTAG
AAATAATGCTCAGCCAAGCGAGGTTGGACATATCAGTTCCAGAGTAGGTC
TCAACCCCGTATATGAGTCCAATGAAGCCTGTCTCATCCAGTTAACGGCC
TTTTGAGCAGAGAATCCTCCCTATTTTCGGAGAGGACGCGTCGAATATAA
AGCAGGTCCAAAGAAGCAAGCAATAGCCAAAAGTTTGAAAGGTTAGTACG
AGCAGCGGCTGGAGGACACTATGGTCGTGCAACGGGGGTAAAGGGTTTCA
CGTATTGTAGCAGAGCACGTCAGAGGGTTATTCGTGACATTCGAGGCCAA
CGAGGCGGTAGGACTTCGTAAGCGCATGACCATCCCGGTCACAAACGTAG
TGCGGAGCGCCTCGTCACGCTCAACAAGGCCCTAGAACGCGCGGCGCAGA
TCGACCCTTTTAAACGCCGGCACCGAGCCGGACCGTCCTGCCCAGGTTGT
AAAGCGCTCCATCGGCCGACTTATGGCACTCGAGCCAAATCGCCCGGTTC
CCCATCGGTCAGCGCAAACGGCCCCCCCGGGCGTCGCCACCCGCGGCGAC
GAGGGGCCGTCCAGACGGGTGATCTCTCTCGTGAGCTCGCGGAGAGAGCC
TCCTCGCAAGATCGATGTCAGCGGGATCGCGCGCCCCGTCCGCACCTGAA
ACGCGTGCTGGAGCTCGACGGCAGCGAGGGAGTCGAGGCCGAACCGCGAT
ATCGGCAGCGCGTCGTCGATCTGCCCGGCGTCCAGACGAAGCGCGCGGGC
GAGGGTCGAGCGCAGCGCGTCCAGCAGGCTCCGGCCGGAGGGCTCCTCGG
TCTCCGGGGGCGCGTCGTCCGGGGGCGAGGCGTCGTCGAGGAGCTCCGGC

GCGAACGCGACGTGGCGCTCGCCGAGCGCGTCCTCGAGAAAGGCGCGCCG
GCACTCCCTCCGGCGGACCTTCCCGCTCGACGTCTTCGGCAGCGCGCCCG
GCGCGATCAGCGCGACGGCGTGCGCGACGAGCTGGTGCTCGGCGGTCACC
GCCTCGCGCACGGCCGCCACGATCTCGCGCGGATCCGCGGCCACGCGCGG
GTCGACCTCGCACACCACGGCGAGGCGCTCCTCGCCCTCGTGCTCCACGG
AGAACGCGGCGCTGCAGCCCGGCCGGACGGCGCGATGGCTGCTCTCGACG
GTCTTCTCGATGTCCTGCGGGAAGTGGTTGCGGCCTCGAAGGATGATGAG
GTCCTTCGACCTCCCCACCACGAACAGCTCGCCGCCCCGGAGGAAGCCGA
GATCTCCCGTGCGCAGGTAGCGCGGCGCCGCGCTGCCAGCGAGCGTGGCC
CCGAACGTGGCCTCCGTCTCCTCCGGGCGCCCCCAGTAGCCGACGGCTAC
GCTGGGCCCGGACACCCAGATCTCCCCGATCTCCCCCGGCCCGAGCTCGT
TCCCCGCGGGATCGACGATCGCGACCGCCCGCGGATCGAGCGCCCGACCG
CTGCCGACGAACACGCGCGCGCCCTCCGCCGCCGACGCGACGGCGCGCCC
GAGCTCCACCTCCTCGGGGGCGAGGCGCGCCAGCACCGGCGCCTCGGCCC
GCGCTCCGCCGCTCACGATGAGCGTGGCCTCGGCGAGCCCGTAGCAGGGA
TAGAACGCCTCTCGCCGGAACCCGCTGACCGCGAAGGCGCGCGCGAAGCG
ATCGAGCGTGTCGGCGCGCACCGGCTCGGCGCCCGTGAACGCGACCTCCC
ACGACCGCAGATCGAGCGCCGCTCGCTCCTCCTCCGAGCTCTTCCGGACG
CACAGGTCGTATGCGAAGTTCGGGCCGCCGCTCACCGAGGCGCCGAGCGC
CGAGACGGCGCGGAGCCACCGCATCGGCCTCTGCAGGAACGAGAGCGGCG
ACATGAGCGCGACGCGGATCCGCCGGTAGAGCGCCTGCAAGATCCCGCCG
ATGAGCCCCATGTCGTGATACGGCGGCAGCCAGATCACCCCGACCGGATC
CGGGCTCGTCAGGTCGAATCCATGCGCGATGAGCCGCGAGTTGTGCAGCA
GATTCCCGTGGGTGAGCATCACCCCCTTGGGCTCGCCGGTCGAGCCGGAG
GTGTATTGAAGGAACGCGACCGACTCCGGCCGGAGCGCCGCGCCCGGCCC
CTCGATCGGGCCCGGCGACGGGCCGTCGGTCGCGATCCACCGGAGCCGCT
GCAGCGCGGCGGCCGCGGCGCTGGCCGGCAGGGACGCCACGATGCCGGCG
ACGGCCGATGACGTGAGCGCCGCCTCGGCGCGCGCGTCCGCGACGATGGA
AGCGACGCGCGGCAGCGTCCGCTCGAGCCGGCCGAGATCCGGCGGATAGG
CGGGCACGGTCCGGACTCCAGCGTAAAGACACCCGAAGAACGCGGTGATG

```
TACTCGATCCCCGGCGGATACAGCAGCAGCGCGCGGCCCCGGGGGCGAC
GCCCGATGCCTGCAAGAGGGCCGCGACGGTTCGCGCGCGCTCGTCAATTT
CCCGCAGGGTCACCCAGGTCGCCCCGGCCTCGACGTCGCCGGACTCAAGA
AAGCAATAGATTGGGCGGGCGGGCTCAGCTTCGGCCCGCTGGCGCAAGAG
GTCGATAACGGTGGAAGGGCGGTTCCGTTCGTTCCGTTCCAATGCAAGAA
AAGCATCATTCATTGAACAGACCCCTCCGCCGCGGAGATAGCAGCTTGTC
CGCTGCGACACAACCGCCGCGCGACGCGCGTGGCACGGCGGGATCCGGGC
GTTACTCCACCTGCACTTCCCGTCGCGTCACGCTCGCTCCGCCGCGGGTG
TCGTGAACCACCGCCCACAGCGACACGCGCCCTGGCTCCGAGGGCGGCGT
CCACGTGGTCCCGTTGCCCCCGCGGGAGGCGCCGGTCGTATCGCTCACCA
GGCGGCGCGCCCCGTCGAACTCGCCACCGTCCGTGTAATAGTCGACCCAG
ATCGCCTCGCGCGCCGGTGGACCGCCGAGCCCGGCGGCTTCCTCGTCCAC
CTCGGCGGCCTTCTCGGGCACGACAGCCTCAATCTCATAGGTCGTGCACT
CGTCCTCGGCCGGCTCGGTCCGGCCGCAGCCTTGGGCCTGCTCCTCGGAC
CGAACGCACCGCTTCACGACGGGCAAACCGTCCTCGCCCGGCGCGACCTC
ATTGCCATCGAGCTTCAGCGTGAAGCCGTCGATGGGCGGGTTCGTGTTCA
GCCGCTCCTTCTTGAAGACATAGACCTGCGTGTAGCCCACGACGAAGCTG
TCCGGACCGAGCACCGTCCCGTCGTCGCCGACGCACTCCAGCGGAAACCC
GGCCGTTTCGGGCGCCGAAGCCACGCGTGTCGTGCCGGCGCACACGGCGA
ACAGCACGTAAGCCGACGAGTACACCGTCCCCGTCTCGGTGGGCCTCGCG
TCCTTGAGGATCTCCTTGGGCAGCTTCCACCCGAACGAGACCGCATCGGG
CTCGCCGCTCTTCTCCGGACCGATCTCCTGCTGCGCGAAGGGGACGGTGC
GCTCCCCGTCGCCGCTGCCGCTGCCGCCGTCGCCGCTGCTGCCGCCGTCG
CCGCTGGCACCGCCACCGCCGCCGCCGCCAGCGCCACCATTGCCGCCTTC
GCCGCCTCCAGCGCCACCACTGCCGCCGTCGCCGCCGCTGCCACCATCGC
CACCACTGCCGCCGCCGCCGCTTCCGCCGCTGCCGGCCGGCACCGCCTCC
CGGATTCGCGACGATTCCCACCGCATAGGTGCCCAGCCACTGCGGGATGC
ACCCGAGGTGCTCGTCCACCCCGACCGGCGGATTCACGCAGCCGCCCACC
CACGTGACCTCGACCTTCCGCGGCGCGCCGCCCTCCGCGCCTTTCGCGTC
GGCGTACGTCATCCGGAACGTCACGAGCTCTTCCGCCGCCGCGTACGGCT
```

TGTCCGCCGTCACGGCGAGGACGCGGAGCCCCTTCACCTCGGACGAAGGG
GCCATGTCGCTCCCGGCGCAGCAAGGGATGCCCACGGCCAGGGTCGAGAG
CGCGGCCAGCAGCGCGCGACGAGCGGGCAGTGCGGTCCGTTTCATCAGAA
ATCTCCTCGCAGCCCGAGCGTGGGCAGGAAGGGGAGCCCCGTCACGTACT
CCCGCTTCGTGTAGTTGAAGTTGTAGCTGATGCCCTCCGCAGCCATGTAA
TTGTAGACGTTCTGGATATCGAGGTAGAGCCCGAGCTGCCACCTCTTGAA
TTTCCACGTCTTGTCGGCGCGGATGTCGAGCTGGTGAAACAGCGGCATCC
GCTCGCTGTAGTCACCCCCGAGCGGGATCGGCGAATACCTCGCCGAGGAC
GCGTGGTAGATCGCGTTCACCCGGTTCGGATTGCACCCCTTCTCCTCCGG
ATCGCAGACATAGGGCGTCTGCAGGTTGCCCGACACGAGCCGGAAGCGCG
CGCCCAGCTCCCAGCCCCGGCCGAGCCGCAGGCTCCCGAGCACCGTCAGC
ACGTGCGTCTGATCGAACTGGGTGAGGTGCTCCTCCTCGTCGGGGCCGTC
CTTGCGCACCGACCGCGAGAGGGTGTACGCCGCCCAGCCGAAGAAGCGCT
CGTCCGGCTTGTACTTCAACAAGAGCTCGCCGCCGACCGCGTATCCGGTG
CCATCGTTGGCATAGTCGTCCTTCTCCGGCGAGAAGACGACCAGCCGATC
GAGCTGCTTGTAGAACCCGTCCAGCGTCACCTCGATCTGCGGCGTGATCT
CCTGCTCCACGCCGAGGCCGTAATGCACGGCGCGGTTCGACTTGAGCTCC
GCATTGCCGAACGGCTCGATGCTCTCCGCGAACTGCGGCGCCTGATAATA
AAGGCCCACGCCCCCCTTGGCCGTCGTCCGCGGGAAGCCGCTCCGGATGT
CGTAGCGCGCGTTGACCCGCGGGCTCACGTCGAGCGTCTGCGTATCGAGC
GCGTAGTCGACCCGCACCCCGGGGACGATCCGCGCCCGCGGCGAGGGGAC
GACCTCGAGCTCGGCATACGCCGCGGGCCGCGAGTACGCGCCGTCGAACG
ACCGATCCTGGAACGGGTACGTCGAGAACGGCTGGTTCGACGGGTGGCCC
GCGGGCTGCTGCGACGGCGCGCGGATGTTGACCGTGGCGACGCCGCCCGA
GAGGTCGGTGCCGACGTTCATCGTGAGGTACCGCGCGAACCTGTGCGAGA
GCTCCAGCCGCAGGTCGAGCGAGGTCGAGACGACGTTGAAGGCGAGGGGA
GAGATCTCGAAGTCGGCGATGTCCCGGCCGAGCGCCATCGACCACAGCAG
CCGATCCCGGCTCCCGATCCGGTTCTCGTAGCTGAGCTGGAAGCGCTGGA
AGGCGGTGTGCAGCCCGAAATCGCCCGTCAGCGCCGGCTCGTCCTCCGGC
GGCTTGTCCAGGGTGATCTTGAAGGCGTCGTCCGATCCGTAGAAGCTCGC

GCGCACGCGCTCGCTCGCGGAGGGGCGGCCCTCGAGGACGAACTGGTAAT
CATAGTAGACGGGCGCCTGCGTGACGCTGGAGCCCGCCTCCTTGAGCACG
GGCCCGAGCCACGCGTCGACCCAGCTGCGGCGGCCCGCCGCGATGAACGT
CCAGTCCTTGAGGAACGGGACGGGGCCCTCGAGGAGCACGCGCCCGTCGA
TGAGGTCGAGCTGGACCACGCCGTGGTACTTGCCGTCCTGCTTCGGCGAG
CGGAGCCCGACGTCGACGATGCCGCCCATGGCGCGGCCGTACACGGCGCT
GAAGTTGCCCGGATAGAAGTCGATCTTCTCGAGCATCTCGGTCGGCACGA
CCGAGGAGAGGCCGCCGAAGTGGTAGATGATCGGCACCGGGGTGCGATCG
ACGAACGTGAGCGTGTCCTGGGGCGCGGACCCGCGCACGATGAGCAGCCC
GAAGCCGCTGCGCGCGACGCCCGGCAGGCTCTGCAGCGACCGCAGCGCGT
CGCCGCCGGTGCCGGGGATGCGGTCGATCTCGCGGCGCTCGATCGTCCTC
CGCGTCACCTCGCGCGGCGGGCGCTCGCCCTGCACGGTCACCTCGATGCC
CGGCGCCTTGCCGTCCTGCGGCGCGGCGAGCGAGATGCGGTAGCGCACCT
CGATCGCCTCGCCGGCCGCGATCTCCTCCTCGGCGGCGAACGGCTCGAAC
CCCGCGGCGGCGACCTCGACGCGGTACTTGCCGGGGGGGAGATTCTTGAA
GCGGAACTTGCCGCCCTGGTCCGTCTTCGCCTCCTCGCGGCCGCCGTCGG
GGCGCACGAGGGTGACCGCGATGTCCGGGAGCGGCTCGCCGGTGCCCGCG
GACAGGACGGTCCCGACCACCGTCTCGACGTCGGCGGGCGGCGCCGGCGC
GGCCGCATCGGCGGGCTTGGGCGTGAGCGTGAACGCGTACCGGTAGAGGA
TGCGCGCCGCCGCGGGCGTGCCGTCCGGGCGCCGCGCCGGCGCGAACTCC
AGGCCGGGCGCGGCCCGCGAGCGCCGCCTCGTTGAAGCCGTGCCCGCCGG
GCGTCGCGACCTCGGCCTTGGTGACGCGCCCGGTCTTGTCGATGTCGAGC
TTGAGGATGACGCTGCCCTCGACGCCGGCGCGCTGGGCCTCGATCGGATA
CGCGGGCGGGGAGTACTTGATCAGCGTCGGCGGGCTGATGGCGGCGGGCG
CCGGCGGGGGCGCGCCCGGCTGAGGGACGACGACCGCGCCGGCGCCGCCG
CGGGGGACCGAGGCGCCGCCCGAGTCGCCCTCGGCGGCAGGAGGAGGCTC
GGGCGGGGGCGCGCCGGCGGGCTGCGCGCGCGCTGCGCTGCCGGTCATCG
CGACCGCGAGCAGCAGCGCTTCCGAGACGACGAGGCGCATCACGGAGGAC
GCTGTGGAAGGCATGCGGCCCGCCCTCTCGCATGGCGAGGCCGAGGCGGA
AAGACGCATCGCGCAGCCAGGACCGTGCTTCACATTGCTTCACACAACGG

GCGCCGCGCGCGCTCCCGGGCCGCGCGAGCGCAGGCGGCGCGCGCGCCCG
CGGGCGGCGCGATCGCGAGCGGCGCGCGGTGCGATCAGCCGCCGACCTCG
GCCACGAACCGGCTCACGTCGTCGCTGTCGCCCACGAGCAGCAGCGTGTC
GCCGTCGCGGATCACGTAGTCCGGTGTGGGCGCCTCGAGCCGCGGCTTGT
CGCCGGGCCGCTTGTTCGTGTGCGGCCGCACACCGAGCACGTTGATGCGG
TACCGCTGGCGGATCTTCGAGCCGGCCAGCGTCTGCCCGACCAGCGGCCC
GTGGGCGTTCCAGGGGACCACGCGGTAGTGGCTCGCGAGGTCGAGGAGGT
CCTGCGCGAGCGGCATGGTGATGTCGGCGCCGACGCGGCGGCCCATCTCG
GTCTCGAGCTGGATGACGCGGGTCGCGCCCACCGCGCGCAGGATGTCGGC
CTGGCGATCGGTGGCGGCGCGCGCGATGATCTCGCGCACGCCCATCCGGA
CGAGGGAGGCCACGCAGAGCACGGACGGCTCGAAGTGCTCGCCGAAGGTC
ACGATCGCGGTCTCCACGTACTGCGCGCCGATCCCCTCGAGCACCTTGTG
GACGGTGGCGTCGCCGACGAACGCGGCCGAGGTCTTGTCCTTCACGGCGT
CGACGGCCTCCGGGTTGTTGTCGACCGCGATCACCTCGGCCCGGTTCTTC
CAGAGGGTCTCGACGACCGACGTGCCGAACCGCCCGAGCCCCGATGACGA
GGACGCTCTTCGATTTCATGGTCTCCGGTCGCGCGCCGCCCCCTCGGGGC
GCGGCGCCGCGAAGGTCTCACGGATGCGCCGGGAGCGCCACGGTTCGGCG
CTGCCGTCGTCGCGACGGCGCGGCCCGCGCGGGCCGCGCCGCGCGGCGCT
CAGTAGAGCTCGTCCTGGTGCTGCCACCGCTCCGAGATCCACGCCTTCAG
GTACGCGATCTCCTCCTCGTACGTCGTGAAGTCGTCGCGCCAGCTCCAGC
CCTCGTAGCTCCGGTACGCCTCGCCCCACCGCGCCTCGTCCCGGCGCGCG
CTCGCGTCGATGCGCTCCACGTAGCCGTCCACGATCGCGTGGATCTCGGC
CTCGGCGAGCGCGCCGCGCAGGACCTGATCGTAGCGGGCGCGCAGCGGGT
CGCCGATCGACGGCTCCTCGAGGAGGCGCTCGAAGAGGAGGTTCACGTCG
CGGTAGTCGACGCGATCCGACGCCGGCTCGCGCTCGGTCTCCCACGACTG
GCCGAAGCTCGCGTTGAAGTCCCACGGCGCGTAGCGGAATACGCCGTCCG
CGGCCGGATCGCGGTAGTGGTAGCTGTTCTTTCCGGCCGAGTCGTTGGCC
ACGATGAACGTGACGAAGATCCACCAGTCCTCGTAGTCGCGCAGATCGAT
CCGCGACCCGATCTCGGCGGCGAACGTGGCGTCGTCGGACTCGGCCACGA
AGCTCACGAGATCTTCCAGATCCGAGAACGCCTCCGGCTCGCCCTCGGCC

GGCGCCCCTTCCTTCTTCTCGAAGCCGTCGTGCAGCGTGTCCTTGGGGTC
GCCGGACCGGTCGGTCAGCGCGAAGTTCGCGTCGTGGCTGACCGCCTTGT
AGAGGTTGCCGTCCTGCGGGTAGCCGTGGTCCTCCATCAGGTAGCCGTCG
ACGTGATCCGCGACGGTGTAGAGCCCCGCGTACTCCCCGTCGAGGTACAG
GACGGCGCTGTAGGTCTTGATCTGGATGTGCTCGGGATCGAGGCGGTTCC
AGAGGTCATAGGCGAGGCGCTGCCGGACATAGGAGTTGTCGTCGAACGTC
GTGATGAGCACGACCTTGCGGCGATCGGTGAAGCCGCCCGCCTCGTCGGG
CTCGTTGAACTTGTCGTCCTTGGGGAACTTGAGGGTGTAGCTCCGCTTCG
GGTACGAGAGCGAGCTCTCGCCGCGGAGCTCCGCCTCCGCGGCGTACGTG
TGGCCGCGGTAGATCACCGTGGCCGGGGCGTACTCCTTGTCCTCGGGGAC
GGGCGAGAGGAAGAGCACCGGCAGGCCGTACTCCTCGGGGTAGCGGGTCG
GATCGACGACGGGCACGTTCGACGGATCGGCGAAGGCGTCGGCGACGCCG
ACCTTGACGCGCCCGACCTCGGACGTCTGCGCGACGCGGATCTCGATGTC
GTAGACGGCGGCCTGATCGAGCCCGGGCGAGAACGTCACCTCGCGCGCGA
TCGGGTCGTACGCGGCGCCCTCGGGGAGCGGGCCGACCTCGAACGCGTCG
CCGGCGAGCGCGAGGCCGCTCGCGCACGTCACCGGGAACGTCACGGTCTC
CCCCTCGAGGAGCCAGTGCGGGCCGCCGCCCGACGGCTGGCAGCGCGAGC
CCTCGGCGCTGGAGCCCCCGCCGCTGGAGCCGGAC

Seq ID No 90 (>Contig11)

GGCGACCCCACATATCACATAGTAGAATCAGTGTGAGTTAGACAAATGTC
GAGTGATGAGAAGGACAGAAGTGAGAACTCTGTCGATCACTGTAGAACGA
GAGAGTATGAGCCTGCATACATGATAGCGGACATGAGAACGAGTGTANTA
TGATGCTACTAAGAGAGTAACAGATCAGAGACTAGAGTAGAGCAATAGAA
NTCAGAGATAAGTCAATGACGAGGAGTAGTGATAGAGCTCTTAATAATGG
CTGAGGTCGAAGATAGAAGTGCATAGAGCGATAGATATACAATCGGTTGA
AGCAGAGAGTAAGATAAGATCAGACACNGAGTACAGAGAGAGACGAATAG
ATGGCGTGATNTCACAGAGAGGTGCGAGCGTAGCTGACGAGAGCAGAGAC
GCAGAGTAAGTCACACCTAGATAGTTACGGCGAGAGACAAATGATAGGAA
GGAGTGGACGAGATCAACAGNCCGGAGCACAAGAACGTGAGATGCGACCG

TGTAATAAACAGGAGACAAGAGCGACTACATAAGAGAGCGAAGCGAATAG
ATAAGATATAAGCCCAGAGCAAAATAGAAGGAGAGAGAGAGTATTTGTAA
TAAAGCAACAAGACGGAGAGAGCGAAGCAGCAGGCAACGATTAGAAGAAA
GACGACAGGAAAGTGAAAGCGAAAGAGAGCAGGTAGAAAGAGAACCAAAA
AAGCACGAAGGAAAAGGAAGCTTCTATGATAGGTGCGGACAAGGCGTAG
CTACAGGAGACAGCCGGCATACGAGGAGCCGGTAAAAGCTAGCCTTTCAG
AACACATCGGGAGCGCGTAAAGGCGGACCACGCTCGACGGGATCATGTAC
GCCGACAGCGACGCCTTCAGCCCCGCGCGCACGTCCGGCGCGTCGCCCGC
CGCGTCGCCGTCGAGCACGACGTAGGCGACCAGGCGGGCGTCGCCCGGCG
CGTCCTCGCGCAGGACCACGGCCGCCTGGCCCACGCCGGGCACGCGCCGG
ATCTGCGCCTCGACGTCGCCGAGCTCGATCCGGTGCCCCCGGAGCTTGAT
CTGGTGGTCCGAGCGGCCCTGGAACTCGAGCATCCCGTCGGGCAAGAAGC
GCGCGACGTCGCCGGTCCGGTACATCCGCCCGCCCGCGGCGCGCGCGCAC
GGGTCGGGCAGGAAGCGCTCCGCGGTGAGCCCGGGCTGCCCCACGTAGCC
GCGCGCGAGCGGCGCGCCCGCGATGTAAAGATCGCCGAGCGCGCCGATGG
CGGGGCGGCGCATCGCGCCGTCGAGCACGAACACCTCGGCGTTCGCGACC
GGCGCGCCGAGGGGGACCCACGTGACCCGCGGGTCGCTCGGCAGGACGCA
GCCGGTCACCGCGATCGCGGCCTCGCTCGGCCCGTACATGTTGATGAGGT
CGCCGTCGTGCTTCGCGTAGAAGCGCCGGACGAGATCGAGCGGCACCGCC
TCGCCGCCCACGAGGACCTTCCGCAGGCTCGCGGGGAACGGCTGCTCGGG
CCCCCCGAGGAACGCCGCGAGCATCGAGGAGACGAAGTACGCGGTCGTCG
CCCCCTCGTCGCGCACGAGGCGCCGAAGGTACTCGGGATCGCGGTGCCCG
CCGGCCCGGGCGACGACGATCCGCGCGCCGAACGAGAGGGGCCAGAAGAT
CTCCCAGACGGAGACGTCGAAGCCGAACGCGGCCTTGAGCAGGACCCGGT
CGTCCGCGGTGAGCGCCCAGTACCGCTGGATCCACTGCATCTGGTTGACG
ATGGCGCGGTGGGAGATGAGGCTCCCCTTCGGCGTGCCCGTCGATCCGGA
CGTGTAGATGACGTACGCGCCGCTGTCCGGCGGCGGGCTCACGGCGGGCC
GCGCGTCGGAGCACGCGGCGATCTCGGCGGCCTCGGCGTCGAGGAGCAGC
GTCGTCCAGCCGCCGGTCGGGAGCTCGTCGGCGATCGCGTCGTGCGTGAC
GAGGAGGCGCGCCCGCGCGTCCCGCATCATGAAGGCGAGGCGCTCGCCGG

GGTACTCGTGGTCGAGCGGCAGGTAGGCGCCGCCCACCTTGAGCACCGCC
AGGGTCGCGACGACCATGTCCTCGGAGCGCGGCACGCAGACCCCGACGAT
CGTGTCGAGCCCGACGCCGCGGCGGCGCAGGCAGCTCGCGAGCCGGTTCG
CGCGCCGCTCGAGCTCGCCGTACGTGAGCGACTTGCCCTCGCTGCGCACC
GCGACGACGTCGGGGTGCTGCTCGGCGCGCTCCTCGAACCACCGGTGCAG
CGCGCAGGCCGACGGCAGCTCCATCGCGGGCCGCGCGACCACGCCTCGA
TCTCGGCGCGCTCGCCGGGGCCGACGTACTCGCCCTGGGCGACCGGACGC
TCGGGGTGGCGCGACAGGTCCTCGAGCAGCGCCGCGAGGCGCTCGGCGAG
GCGCTCGGCGTCGCGGCGCGCGCCGGCCGACGCGTCGTAGCGGAGCTCCA
GCGACGCCGACGGGCCCGCGCCGGCGCAGTGCAGCCGGCAGGCGACCTGG
TCAGACGTGCTCCAGACGTCGAGCACGCGGGCCCGCGCGCCGTCGAGCGA
CAGCGCCGGCGCGGCCTCCGCGGCGGAGAAGCCCCCAGCTCATCCGGTGG
CTCACCCCGGGCGCGGCGTCCTGGTGCGCGGCCGCCTCGGCCTCGGCGAG
CGCGAGCCGCCGCGCGACGTCGGCGAGCGTGTCCGAGGCCGAGATCTCGA
TCCGCACCGGCAGGAACCGCGCGAACGGCCCCACCGCGCCCGCGAGCGCG
TCCAGCGACCGCCCGTCGAAGCGGACGGCCACGGTGACCTCGGGCTCGTT
GCCGCCGCTCATCCGCCACAGGAGCGACGCCCACAGGGCCAGGAGCACGA
TCCGCTGCGGGACCTGCCACGACGACGACCAGCGCTCGACCTGCGCCATC
CCGCCTTGTCCCAGATCGACCCGCGCGCGCCCCGAGCCGGCGCCGGCGCC
GGCGCCGCCGCGGCTGAAGGCGAGGTGGAGCGGGGGCCCGAAATGCGAGC
GGCGCTCGGCCCAGAACCTGCGCCCGTCGCCGGCGTCCTCCGACTCGAGC
ATCCCGTTGAGCCACTCGGCGACGTCCGCGTACTGCTGCTCGGGCGGCGC
GCCCGCGCCCGCGGTCGACGCGCAGAGCTCGCGGACGAGCGGGGCGATCG
ACTCCTCGTCGACGCACCACGCGGGCGCCGCGAGCACGAGCCGGCGCTCC
TCCGGGCCGACGCGGACCAGGCCGACGCGCAGCCCGTCGTCCGCGCCGCG
GTCCTCCGAGAGGCGCGCGACGAGCCGCGACATCCGCTCGCCCTGCTCGG
CTTCGGAGCACCCGACCCAGTCGTCCTGCTGACGCCCACGCGAAGCGCGG
CTCGCCGACCACCTGCGCGGCCTCGCCCGCCCCGCCTCGACGAGGCGCGT
GCGCAAGATCTCGTGCCGCTCGGCCAGCGCGAGCGCCGCCGCCGAGAGCC
GTCCCTCGTCGCACGGGCCGGTCACGGCGACGACGGCCAGCGTCCGGCAC

CCGGGCGCCCCCGCCTCCCGGTCGAGCGCGGATCGCCCGCTGCTGCGG
CGAGAGGCTGAAGCCGGTCATCTCGTGGTCACTCATCCAGGTCGTCCTTC
GGTGAGGTCTTCGCTTCGCCCGGCGCGCCCGGGCGCGGAGGGTCACGGC
GCGCCGGCGCGCGGAGGTCAGCTTGTCGAGCGCGGCGCCGCGCGCGGCCT
TCCGCTCGAGCTCCCGGCGGGCCGCCGCGGCGCGCTCGAGCTCCCCCCGG
AGCTCCGAGACCGGGGTGTCCGGCCGCGCCGTCGCAGTGGCCAGGATCTG
CCGGTAATCGCTAAGGAAATTGTCGACCGTCGCCGCCCGGTACAGCTCGC
TGCTGTGCTCGACGCCGAAGCGGAACGAGCCGCCGGCCTCGGCGACCGTG
AGGACGAAGTCGAACGCCGTCGTGGTCGCCTCGCCCTCCAGCGCCTCGAG
CTCGAGCCCCTCGAGCTTCATCGGGGGGACGTGCACGTTGCGCATGACGA
ACTTCGCGTCGAAGAGGGGCACGTGCCCGACGGCCCCCTTCGGCCGCAGG
GCCTCGACGAGCCGGTCGAACGGCAGGTCCTGGTGCTCGAACGCCTCGAG
CGCGACGTCGCGCACGCGGCGGACCAGCGCGCCGAACGTCGGGTCGCCCC
CGCAGTCGGTCCGGAGCACGAGCTGGTTGACGAAGAAGCCGATCATCGGC
TCGGTCTCGACGCGGTTCCGGTTCGCGACGTCGGTGCCCACGACGAGGTC
CTCGAGCCCGGTGCGCTGGTGCAGGACGAGCTTGTACGCGGCGAGCAGGG
CCATGAAGGGGGAGATCGCCTCCCGCTCGCAGAACGCCTTGATCTGGCGG
GTGAGCTCGGCCCCGGCGTCGAGGCTCCGCCGCGCCCCGCGCCACGTCCT
TCGCCCCGCCGGCTCGTGGTCGACCGGCACGCGGGCCCGGCGCAGCGCGC
CCGAGAGCTTCGTCGTCCAGTACCGGAGCTCGCCCTCCAGGACCTCGCCG
GACAGCCACGCCCGCTGGGCTGCGGCGAAGTCGACGTACTGCGCCGGGAG
CTCCGGCAGCCGGGAAGGCTGGCCCTGCGCGAAGCCGCCGTAGAGCGCGG
CGAGCTCGCCGACGAAGACGCCGACCGACCAGACGTCGAACACGACGTGG
TGCACGACGAGCGCGATGACGTGCTCGTCGTGGCGCTTCCGGATGACCCG
CACGCGGAGGAGCGGCCCGCGGCTCAGGTCGAACGGCGCGAGGCTCTCCT
CGAGGACGAGCGCCGAGACCGCCGCGTCGAGGGCCTCGCCCGCGAGGTGC
TCGAGGTCGGACATCCGGAACGGCACCCGGGCCTCGGGCGCGACGACCGG
GAACGGCACGCCGTCCCTGGCGCTGAACGTCGTCCGCAGCGCCTCGTGGC
GCCGCGCGATCTCGAACAGGCTGCGGCGGAGCGCGTCGACGTCGAGCCGG
CCCGTCGCGCGCACCACGAACGGGATGTTGTACGCCGGGCTGCCCGGCTC

```
GAGCTGATCGACGAACCACAGCCGGTGCTGCGCGAACGACAGCGGGAGCG
GGCCGTCGCGGGGGATCCGCGCGATCGGGGGGAACTCGCGCCGCCGCGCC
TCGCCGCGGCGCGCGGCGTCGACCTGGCCGCGAGCGCCGCGACCGTCGG
CCCCTGGAAGAGCGCCCGCAGCGGGAGCTCGACGCCGAGCTGCGCGCGGA
TCCGGGACATCACCTGGGTCGCGACGAGCGAGTCGCCGTGCAGGCCGAAG
AAGTCGTCGTGGACGCCGATCTCGTGGACGCCGAGGAGGGCGCTCCCAGA
TCGCGGGCGATCGCGCGCTCGGACTCGGTCGACGGCGCGGCGAACGCGGC
CCCGGCGTGCGCGCGGGAGACCGCCGACGTCGGGAGGGCGTTCGGCGCGG
GCGCGAGGGGCGCGTCGGCCGGCGCCGGCTCGGCGGCGGGGGCGCCGCGG
CGCGGCGCGATCCAGTGCCGCGCCCGCTCGAACGGGTACGTCGGCAAGCG
GACGAGCGCGCCGGGGGACGCCCCCGCGGACGGCCGTCCAGTCGACGGC
GTGGCCCGCCTCCCAGAGCTGGCCGAGGGCCTCGGCCAGGCTCGCGGGCT
CGGACGCGGCGTGGGTCGACCCGAGGCTCGCGATCGCGGCGCCGCCGCGC
CCGGCCAGCGTCTGCCGCACCAGCGTGGTCAGCCCGCGGCCGGGGCCGAC
CTCGAGGAACAGGGCGTGCCCGGACGCGAAGAGCGCCTCGACGCCGTCGC
TGAAGCGGACCGGCTGGCGGAGGTGCCGCGCCCAGTAGGCCGGATCGGTC
GCCTCGGCGTCGGTGAGGAGGGCGCCGGTGACGTTCGAGACCACGGGGAT
CTCCGGCGGGGAGAGCCGCGCGCGCCGCACGCTCTCGAGGAACGGGGCCA
CCGCGCCGTCGATGAGCGCGCAGTGGAACGCGTGGGACGTCTGCAGCGGC
CGGGCGAACACCTCGCGCGCCTCGAGGCGCGCGGCGAGATCGCGGATCGC
GCTCGCCGGGCCCGCGACAACCGTGAGCTTCGGGCTGTTGACCGCGGCGA
TCTCCAGGCCGGCCTCGAGGAGGCCCTCGACGTCCGCGGCCGGCAGGCCG
ACGGCCAGCATGCTCCCGGCCGGCGCCGCCTGCATGAAGCGCCCCCGATC
GATGACCAGGGACATCGCGTCCTCGAGCGTGAACACGCCCGCGACGCAGG
CCGCCACGAGCTCGCCGAGGCTGTGGCCGATCATCGCCGCGGGCTCGATC
CCCCAGCTCATCCAGAGCCTGGCGAGCGCGAGCTCGACGGCGAAGAGCGC
GGGCTGCGCCAGCGCGGTGCCGAGCAGCGTGCGCCCGTCGCCCTCGCCCT
CGCGGAAGACGACCTCGCCGAGATCGAGGCCGCGCGCCCGCGCCGCCGCC
GCGCACGCGTTGAAGGCGCTCCGGAACGCCGCCTCCTGCGCGTAGAGCGC
GCGGGCCATCCCGACGGCCTGCGCGCCCTGGCCCGGGAACGCGAAGACGG
```

GCGCGGCTCATCGGGGCGCGCGAGCGCGCTCGCCCCCTCGCGGGCGAGCC
CCTGGATCGCCTCGGCGCGCGTCCGGGCGACGACCGCCCGGCGGTACGGG
TGCTCCGCGCGCCCGGTCTGGAGGGTGAACGCGACGTCGTCGAGCGGGAC
GTCGGTCGCCTCGAGGTGCGCGGCGAGCTGCGCGCAGGCCGTCGACAGCG
CCTCCGGCGTGCGCGCCGAGAGCGTCAGCACGTGATCGCGCTCCGGGGCC
GGGGCGCGGGGCGGCAGCGGGGGCGGCTCCTCGAGCACGACGTGCGCGTT
CGTCCCGCCGATCCCGAACGAGCTCACGCCCGCGCGGCGCGGGCGGAGCT
CGCGGGGCCAGGGCGCCGCCTCCCGCGGACGAAGAACGGGCTCGCCGCG
AGGTCGAGCTTGGGGTTCGGCGCCTCGAAATGGACGCAGGGCGGGATCTC
GCCGCTCCGCACGACGTGCGCCGCCTTGATGAGGCCCGCGACGCCCGCCG
CGGCGTCGAGGTGGCCGATGTTCGCCTTGATCGAGCCGAGCGCGCAGTAC
GCCTTCCTCGGGGTCTTGCGGCGGAAGGCCTGCGTGAGCGCCTCGACCTC
GATCGGATCGCCGATCGCGGTCGCGGTGCCGTGGGCCTCGACGTAGCCGA
TCGAGCCGGGATCGACGCCGGCGACCGACTGCGCCTCGGAGATCGCCGCC
GCCTGGCCGTCGACGCTGGGCGCCATGAAGCCGACCTTGCGCCCGCCGTC
GTTGTTGACGGCGGAGCCCCGGATCACCGCGTGGACCGTGTTTCGGTCGC
GGAGGGCGTCCGCGAGGCGCTTCAGCGCGACGATGCCGACGCCGCTGCCG
CCCACGGTCCCCTCGGCGCGCGCGTCGAACGGCCGGCAGCGGCCGTCGGG
GGAGCAGATGCTGCCGGGCACGTACGGATACCCGCGCTTCTGCGGGATGC
CGATGGAGACGCCGCCCGCCAGCGCGAGATCGCACTGGCCGCCGAGGAGG
CTCTCGCACGCCATGTGGACGGCCACCAGCGACGTCGAGCACGCGGTCTG
CACGACCACGCTCGGCCCGTGGAGGTCGAGTTTGTACGAGACCCGCGTCG
CGAGGTAATCCTTCTCGCTCGCCAGCATGAGCGCGTGCGGATCGACGGTG
GCCGCGAGATCCGGGTGCGAGAGGAGCTGGAGGAGGTACGTGTTGGAGCC
GCACCCCCCGAAGACGCCGATCGCGCCCGGGAACCGGGCCGGATCGCAGC
CGGCGTCCTCCAGGGCGGCGACCGCGCACTCCAGGAAGAGGCGCTGCTGC
GGGTCCATGAGCTGCGCCTCGCGCGGCGAGTACCCGAAATAGGACGCGTC
GAAGCGGTCGATGTCGTCGAGCAGGCCGCCCGCGCAGACGACGGGCGCCC
CGGGGGCCGCGCTCGCGCCGACCGGCGGCTCCTCGCGCTCGCTCTCCGGG
AAGCGCGCGATCGACTCGACGCCGCGCCGCACGTTCTCCCAGAGGGCGTC

GACGCTCGGGGCGCCGGGGAAGCGGCCCGCCATGCCGACGATCGCGATGT
CGCTCCCCCCGTCCTCGGTCTCGATCGGCTCTGACATGGCTATCCTCGCC
CCCGGCGGCGTCGCGCGTCGCGGCGCGCCTCGGCGCGCTGCGCCCCGACG
TCGGCCGGCTCGGCCTTGACCGTCGCCGCGTCGAGCCGCTGCGCCAGTTG
CTCGATGGTCGGGTACTGGAACAGGTCGGTCAGCGACACGGCCTGCGCCG
CGGCGCCCTCGTCGGGCGCGCGCGCCGCGATGCGCTCGGCGAGCAGGCGC
TGCGCGCGCACGAGGAGCAGCGAGGTGAAGCCGAGCTCGAAGAGGTTGTC
GGTCACGCCGACGGCCTCGACCTGCAAGACCTCCGCGAGCACCGAGGCGA
TGAGCCGCTCGGTCGCGGTCCGCGGGGCGACGGCCGCGGCGCGCGGCGCG
ACCGCGGCGGGATCCGGCAGGGCGGCGCGGTCCACCTTGCCGTTCGCGCT
CAGCGGCAGCGCCGGGAGGACGACGACCTCCGCGGGGATCATGTACTCCG
GCAGCTTCTTCCGGACGAAGTCGCGGAGCGCGGCGCCATCGCCGTCGGCG
CCGACGACGTACGCGACCAGGCGCTTCTCGCCCGACGGATCGGTCTTCGC
CGCCACGACCGCCTGCTCGACCGAGGGGTGCTGCGCGAGGGCGGCCTCGA
TCTCGCCGAGCTCGATGCGGAAGCCGCGGATCTTCACCTGATGGTCGGTG
CGCCCGAGCAGCTCGATGGTCCCGTCGGCGAAGTAGCGGCCCAGGTCGCC
TGTCCTGTACAGCCGCTCGCCGGTCGTGGGGTGCTTCAGGAACCGCTCCC
GGGTCCGCGCCTCGTCGCGCCAGTATCCGAGCGCGACGCCGATCCCGCCG
ATGTGGATCTCGCCGGGGACCCCGATCGGACACGGCTCCAGCCCCTCGTC
GAGCACGTAGGTGTGCTGGTTCGCGAGCGGGCGGCCGTAGGGGATGCTGC
GCCACGCCGGGTCGACGTCCGCGATCGGGTGGGCGATCGACCAGATCGAC
GCCTCGGTCGCGCCGCCGAGGCTCACGACGCGGGGCGCGCGGCAGGCCGC
GCGGATGCGATCGGGGAGCTTCAGCGGGATCCAGTCGCCGCTCATCATGA
CGAGGCGGAGCGACGACAGCGCCGGGTCGCCCGCGCCGGGGGACGCGTCC
ATGAGCATCTCCATCAGCGCCGGGACCGAGTTCCACACGGTCACCCGCTC
GCGCTCCACGAGCTCGCGCCAGTGCCCCGGATCCGAGGCGCGGGTACGGT
CGGGGATCACGACGGCGCCTCCGGCGGCGAGCGTCCCGAACACGTCGTAG
ACCGACAGGTCGAAGCTCAGCGACGAGAGCGCGAGCACCCGGTCCTCCGG
GCCGACGTCGAAGCGGCGGTTGATGTCGAGGACCGTGTTCACCGCGCCGC
GGTGGTCGATCATCACGCCCTTGGGCAGCCCCGTGGACCCGGACGTGTAG

ATCACGTAGGCCAGGTCGTCCGTGCTTCCGCCGGGCGGCCGGCGCGCGAC
GGGCTGCTCGCGCCACCGCTCGTCCGCGTCGACGGCGAGGCGCTCGATGC
CCGCGGGCCAGGCGATCGTCCCGTCGACCGCCGACTGCGTGAGGACGAGG
CGGACCTCGGCGTGCTCCAGGAGGTGCCTGAGGCGCTCCTCGGGGAGGCG
AGGGTCCAGGGGCAGGTAGGCGGCGCCGGCGCGCAGCACGCCGAGCACGG
CGGCCACCTGCTCCCAGCCCTTCTCCATGACCACGGCGACGAGCGCGTTC
GCGGTCGCTCCGGAGCGCGAGGCCGCCGCGGCGATCGCCTCGGCGCGCCG
GGCGAGCTCCCCGTAGGTGAGGCGCCGCTCGGCGTCGACGACCGCGCACG
CGTCGGGCTGCTCGACGGCGCGCTCAAAGAACGGCTCCTCCAGCCGGAGG
TGATCCGGGGTTGCGACCGCGGTGTCGTTCCACGCGACGAGGGCGCGCTC
GCGGTCTTCCGGCGCGACGGAGAGCGCGCGGACGCGCTGCGCGGGGTCCT
GAGTGGCGCGCGAGAGCACGCTCTGCATCGTGGCGAGCATCCGGTCGATG
GTCGCCGCGTCGAAGAGGTCGACGTTGTACTGGAGCGAGATCACGTCGCG
CCCGCCGCGCGGCTCGACGCTGAAGCGCAGGTCGAAGCGCGTGGCCTCGA
CCGGGAGATCGAGCGGCTCGATCCGCACCTCGCCGAGCTCGAGCGCCTCG
GTTGGGGCGTTCTGCACGACGAGCATGACCTGGAACAGCGGCGAGCGGCT
CAGGTCGCGGCGGGGGTTGACCGCCTCGACCACCTTCTCGAACGGGGCGT
CCTGGTGCTCGAACGCCTCGAGCGCGACCTTCCGCGCCCGCGAGAGGAGC
TCCTCGAAGGTCGGGTCGCCGCCGAGGTCGAGGCGCATGACGATCGTGTT
CACGAAGAAGCCGACGAGGGGCTCGAGCTCGGGGCGAGGCCGGTTGGCGA
CCGCGGTCCCGATGGCGAGGTCGTCCTGGCCCGAGCTGCGCCGGAGGAGC
ACGCCGAGGGCGGCGAGCAGGACCATGAAGCGGGTGGCGCCGCGGCTCCG
GGCGAGCTCGTCGAGCTGCGCCACGAGGCGCGCGTCGAGCGGGAGGACCC
GCTCCGCGCCGCGGAACGTCTGGACGGGCGGCCGCGGTCGATCGGTCTGG
AGCTCCAGGACCGGCAGCCCGCGGAGGGTCGCTGTCCAGTGAGCGAGCTT
GTCGGCGAGCCGCTTCCCCGCGAGGTGGCGGCGCTGCCACACCGCGAAAT
CGACGTACTGGAGCGGCAGCTCGGGCATGTCCGCGGGCCCGCCGCCCCGC
GCGCGCCGGTAGAGCTCCGCGAGATCGCGGACGAGGGGTTGGAAGGACCA
GGCGTCCGTGACGATGTGGTGCGTGGACAGGACCAGGACGCAGACGTCGT
GGTCGAGGCGGAACAGCCTGGCGCGGAACACGGGCCCGCGCGCGAGGTCG

AACCCCGTGGCCTGCTCGCGCGACGCCCAGGCGCGCGCCGCGGCCTCCGC
CTCGTCCGGGGGCGTGCCGCGGAGGTCGACCACCTCCGCGGGGGCCGCCT
CGGGCTCGCAGATCTTCTGCGCCGGCGTGGGGCTCGCGACGAACACCGTG
CGCAGGCTCCAGTGCCGCCGGACGAGCGCGGCGAGCGCGGAGGACAGCGC
GTCGACGTCGACGAGGTTCCGCAGGCGGACCGCCTGCACCACGTTGTAGG
CGGTCCCGCCGGGGAGCAGCTGCTCGAGGACCCACAGGCGCTCTTGCTCG
TACGAGAGCGGATACGGCTCGTCCGCCGGCGCGCGGCCCAGCGAGGGCGC
GATCTCGCTGGCGGGCACCGTCGCTGCGGCGGCGGTGGTCGAGGCGGCGC
CGGAGGAGAGGCGATCGGCGAGCTGGTGGAGGGTTGGGTGCTCGAAGAGC
GTGCGGAGGGTGGTGCGGATGCCGAGGGAGGACTCGATGCGTCCGAGGAC
CTGCATGGCGAGCAGGGAGTGGCCGCCGAGGTCGAAGAAGCTGTCGTGTC
GTCCGACGCGGTCGAGGTGGAGGACGGATTGCCAGATGTGCGCGAGCTCC
CGCTCGAGCTCGCCCGAAGGGGGCTCGTAGTCGGCGTGCGCGGCGGGTGG
CGCAGGGAGGAGCTTCTTGTCGACCTTGCCCGAGAGGGACATGGGCAAGG
CGGGGAGCAGGACGAAGTGGGCGGGCACCAAGGCGTCGGGCACCAGGCGG
GCCATGCCCTCGCGCAGGTCGCGCTCGGAGGGCGGGTCGGCGCCCGGCAC
GACATAGGCAATCAGGCGCGCGGCGCTGCCTTGGCCGTGGAGGACGACGA
CGCCCTCGCGGACGGCGGGCAAGCGTCGCAGGGCGGATTCGACCTCGCCG
AGCTCGACACGGCGACCGCGGAGCTTGACCTGCTCGTCGCGGCGTCCGGC
GAAGGCGAGCTGTCCGTCGGGGCGCCAGCGCACCAGGTCGCCGGTGCGGT
AGAGGCGTGCGCCGGGCTGGCCGAAGGGATCGGGCAGGAAGCGCTCTGCG
GTCAGGTCCGTGCGTGTGTAGCCCTGGGCGAGGCACGCTCCGCCGATGTA
CAGCTCGCCGAGGACGCCGGGCGGACGGGCTGCATGTGCGGGTCGAGGA
CGTAGACGAGGGCGCTGTCGATGGGTCGGCCGAGCGGGGGCTCGTCGCCG
AGGTCGGCGACCTCGGCGACGGTGGTGATGACGGTGGCCTCGGTGGGGCC
GTACATGTTGAAGAGGCGGAAAGGAGCGGTCGCCGGAGCGGATGGAGCT
TGTCGCCGCCGACGGTCATCGCGCGCAGGGCGATGCCGGTCCAGTCTTGC
TCGAAGCACGCCTCGGCCAGGGGCGTGGGCATGAACGAGAGCGTGGCCCG
CTGAGCGACAAGCCAGGAGACGAGCGCTGTGGGAGAGCGGAGCGCGTCGT
CGTCGGCGAGGAGGAGTGCAGCGCCGCAGGCGAGCGGCGTCCAGATCTCG

TAGACGGAGGCGTCGAAGCCGCTGGAGGCCAGCTGAGTCCAGCGATCGCG
GGGTGAGAGCGCGAGCAGGTGCTGGAAGAAGGAGACGAGCCTTGAAAGGC
TCGCATGGCGCACACAGACGCCCTTCGGCGTGCCGGAAGAGCCGGAGGTG
AAGAGGACATAGGCCAGGTCGTCGGGCCTGGAGACGAGAGGAATGTGGGT
GCTGGGCGCGCACGCCCCGTCCTGGACGAGGTGGACGGGGCAGGGGGCGG
CGGTGAGCTTGTGGCTGGCCTGGCTGCTGGTGAGCACGAGCGCGGCGCGG
CAGTCGGCGAGCATCTCGGCCAGGCGCGCCGGGGGGTTGGCGGGGTCGAG
CGAGGCATAGGCGGCGCCTGCCTTGAGGACGGCGAGCTGGGCGGCGACCA
TGCGGGGCGAGCGCTCGATGCAGACGCCGACGACGCTGCCGGGGCCGACG
CCGCGGTCGCGCAGCCACAGGGCGAGCTCGGTGGACCAGGTGCTGAGCTC
TGCGTAGGTGAAGCGCTGGTGTCCGAACTCGAGCGCCGTGGCGTCCGGCT
GTCGAGCGGCGTGGGCCTCGAAGAGCGCATGGACGCAGGCGGGGGCCGGG
GCGGAGGCGGCCTGTCGTGCGGCGGCAGCGCCGCTCCAGTCGTCGAGGAG
CAATGCGCGCTCGGCGTCGGAGAGCATCCGGAGCTCGGAGAGCGGTCGAC
CGGGGTGCTCGACGGCGCTTTCGAGCAGGAGCACGAAGTGGCGCGCCATC
CGCTCGATGGTGGCGGGGTCGAAGAGCTGCTGGTCGTACTCGAAGCGCAG
GGCGATGCCGGAGTCGAGCTCTGCGGCGAACAAGGCGAGATCGAACTCGG
CCGCTGCCTGCTCGTCGGCGAGCGTGGTGAGCTCGAGCTCTCCCTGCGCG
ATCCGCACGTCCTCCGCGCCGGTGGTGAGCGCGGCGAGGCGGGGATCCAG
CGACGGCAGAGCGCCCTGGAAGGCGAAGGCGACGTCGAAGAGCGCGCCGC
CTCGCCGGGCCGCGCCCCGGGGCTCTGCGAGCAGGTGCTGGAGGGCGCTG
TCGCCGTGGGCCAGCCCGTCGAGGAACGCGTCGCGCACGCGGGCGACGAG
CGCGTCGAAGGACGCGGCCCCGCGCAGCGCCACGCGCACGGGGAGCATCT
GGACGAAATAGCCGAAAGCCCGAGTGCTCTCGTCGTCGTTCCGCCCCGCC
GAGGGGACGCCCACGACAAGGTCGTTCTGCCCGCTCGCGCGATGGAGCAA
GACGGTGAGCGCCGACAGCAGGACCGAGAAGAGCGTGGTCCCGCGCTCGC
GCGCGAGGCGCGCCAGCGCTCCGGTCAGGGGCTTTGGCAGCGTGATCGCG
TGAGCGCGACCGCGGCGAGGGCTCGCGTCGTGGCGGGCCCGGTCGCGGGG
AAGGTCGATGGCGGTCGTCGCGCCGTCGAGCGCCTTGCGCCAGTATTCTG
CTCCGCCGGCCGCCTCCCGCGGCGAGGGACAGCTCACGCCGGCGGCGAAG

AAGCTCGACGGCGGCGGCAGCTGCGGGGGCCGGCCCGCGCAGCGCCGA
GTACAGCTCCCCCAGCTCGCGAACGAGCAGCGCGAACGACCAGTAGTCGA
CCACGAGGTGGTGAACGACCACCGTGAGCAGCGGCGGCTGCCCCTCTCCG
CGCCGCCAGACATGCACCCGGAGCAGCGGTCCGCGCTCCAGGTCGAACGC
GCGGCGGCGCACCTCGTCCGCGCGGGCGACGATCTCGCGCTCGTCCAGCG
CCATCGCCGGCTCTTCGGCCCATTCCAGGGCGACATGGCGGTGGACCTGC
TGCAGCGGATGGCCGTCGCGCGTGAGGAACGTCGTGCGGAGCGCCTCGTG
CCGCTCGACGAGGCCCTCGAACGCGCGGCGCAGCGCGGCCACGTCGACGC
CGGCACCGAGCCGGACCGTCCTGCCCAGGTTGTAGAGCGCGCCGTCGGCC
GACTTCTGGCACTCCAGCCACATCGCCCGCTGCCCCTCGGTCAGCGCAAA
CGGCTCTTCCGGCGTCGCCACCCGCGGCGACGAGGGGCCGTCCGGACGGG
TGAGCCCTCGCTCCAGGGCCGTCGCTGCGGCGGCGGAGGTCGAGGCGGCG
CCGGAGGAGAGATGACTGGCGAGCTGCGCGAGGGTTGGGTGCTCGAAGAG
CGTGCGGAGGGTGGTGCGGATGCCGAGGGAGGACTCGATGCGTCCGAGGA
CCTGCATGGCGAGCAGGGAGTGGCCGCCGAGGTCGAAGAAGCTGTCGTGT
CGTCCGACGCGGTCGAGGTGGAGGACGGATTGCCAGATGTGGGCGAGCTC
GAGCTCGAGCTCGCCCGAGGGGGGCTCGTAGTCGGCGTGCGCGGCGGGGG
GCGCAGGGAGGAGCTTCTTGTCGACCTTGCCCGAGAGGGACATGGGCAAG
GCGGGGAGCAGGACGAAGTGGGCGGGCACCAGGGCGTCGGGCACCAGGCG
GGCCATGCCTTCGCGCAGGTCGCGCTCGGAGGGCGGGTGGGCGTCTGGCA
CGACATGGGCAATCAGGTGCGCGGCGCTGCCTTGGCCGTGGAGGACGACG
ATGCCCTCGCGGACGCCGGGCAAGCGTCGCAGGACGGATTCGACCTCGCC
GAGCTCGACGCGGCGACCGCGGAGCTTGACCTGCTCGTCGCGGCGCCCCG
CGAAGGCGAGCTGTCCGTCGGGGCGCCAGCGCACCAGGTCGCCGGTGCGG
TAGAGGCGTGCGCCGGGCTGGCCGAAGGGATCGGGCAGGAAGCGCTCTGC
GGTCAGGTCCGTGCGTGTGTAGCCCTGGGCGAGGCACGCTCCGCCGATGT
ACAGCTCGCCGAGGGCGCCGGGCGGGACGGCTGCATGTGCGGGTCGAGG
ACGTAGACGAGGGCGCTGTCGACGGGTCGGCCGAGCGGGGCTCGGCGCC
GAGGTCGGCGATCTCGGCGACCGTGGTGATGACGGTGGCCTCGGTGGGCC
CGTACATGTTGAAGAGGCGGAAAGGGAGCGGGCGCCGGAGCGGATGGAGC

TTGTCGCCGCCGACGGTCATCGCGCGCAGGGCGGAGCCGGTCCAGTCTTG
CTCGAAGCACGCCTCGGCCAGGGGCGTGGGCATGAATGAGAGTGTGGCCC
GCTGAGCGACAAGCCATGAGACGAGCGCCGTGGGAGAGCGGAGCGCGTCG
TCGTCGGCGAGGAGGAGGGCAGCGCCGCAGGCGAGCGGCGTCCAGATCTC
GTAGACGGAGGCGTCGAAGCCGCTGGAGGCAACCTGAGTCCAGCGGTCGC
TGGGCGAGAGATCGAGTCGGAGGTGGAGGAAGGAGACGAGCCTTGAAAGG
CTCGCATGGCGCACACAGACGCCCTTGGGGGTGCCGGTGGAGCCGGAGGT
GAAGAGGACATAGGCCAGGTCGTCGGGCCTGGAGACGAGAGGAATGTGGG
TGCTGGGCGCGCACGCCCCGTCCTGGACGAGGTGGACGGGGCAGGGGGCG
GCGGTGAGCTTGTGGCTGGCCTGGCTGCTGGTGAGCGCGAGCGAGGCGCG
GCAGTCGGCGAGCATCTCGGCCAGGCGTGCCGGGGGGTTGGCGGGGTCGA
GCGAGGCATAGGCGGCGCCTGCCTTGAGGACGGCGAGCTGGGCGGCGACC
ATGCGGGGCGAGCGCTCGATGCAGACGCCGACGACGCTGCCGGGGCCGAC
GCCGCGGTCGCGCAGCCACAGGGCGAGCTCGGTGGACCAGGTGCTGAGCT
GTGCGTAGGTGAAGCGCTGGTGGCCGAACTCGAGCGCGGTGGCGTCCGGC
TGTCGAGCGGCGTGGGCCTCGAACAGCGCGTGGACGCAGGCGGGGGCCGG
GGCGGAGGCGGCCTGTCGTGCGGCGGCAGCGCCGCTCCAGTCGTCGAGGA
GCAATGCGCGCTCGGCGTCGGAGAGCATCCGGAGCTCGGAGAGCGGTCGA
CCGGGGTGCTCGACGGCGCTTTCGAGCAGGACCACGAAGTGGCGCGCCAT
CCGCTCGATGGTGGCGGGGTCGAAGAGCTGCTGGTCGTACTCGAAGCGCA
GGGCGATGCCGGCGTCGAGCTCTGCGGCGAACAAGGCGAGATCGAACTCG
GCCGCTGCCTGCTCGTCGGCGAGCGTGGTGAGCTCGAGCTCTCCCTGCGC
GATCCGCACGTCCCCCACGCCGATCGCGAGGGCTGACAGGCGTGCATCCA
GCGATGGCGGGGTGCTCTGGAAGGCGAAGGCGACGTCGAACAGCGCGTCT
CGCTGCGCCTCGCCCTGCGCTCGCGCGAGCAGGTGCCGGAGGGCGCTGTC
GCCGTGGGCCAGCGCGTCGAGGAACGCATCTCGCACGCGGGCGACGAGCG
CGTCGAAGGACGCGGCCCCGCGCAGCGCCACGCGCACGGGGAGCATCTGG
ACGAAGTAGCCAAAGGCCCTGGCGCTCTCGTCGTCGTGCCGCCCCGCCGA
GGGGACGCCCACGACAAGGTCGCTCTGTCCGCTGGCGCGATGGAGCAAGA
CGGTGAGCGCCGACAGCAGGACCGAGAAGAGCGTGGTCCCGCGCTCGCGC

GCGAGGCGCGCCAGCGCTCCGGTCAGGGGCTTTGGCAGCGTGATCGCGTG
AGCGCGACCGCGGCGAGCGCCCGCGTCGTGGCGAGCCCGGTCGCGGGGGA
GGTCGATGGCGGTCGTCGCGCCGTCGAGCGCCTTGCGCCAGTATTCTGCT
CCGCCGGCCGCCTCCCGCGGCGAGGGACAGCTCACGCCGGCGGCGAAGAA
GCTCGACGGCGGCGGCAGCTGCGGGGGCCGGCCCGCGCGCAGCGCCGAGT
ACAGCTCCCCCAGCTCGCGAACGATCAGTGCGAACGACCAGTATTCGACC
ACTACGTGGTGATCCACCACCGTCAGCACTCTGCGTTTAGTCCTTTCTCC
TGTTCGGCCTATTAATTGCTACTATGGATCCACACTGCTCCGCCTCTTGT
ATCTCCCTTATCTGCACTTGCTGCGCTTNACCTTTACGTTCCTCTCCCTG
CTCTATACTATTTCTTCCCCCGCTTCTCGTCCTATTCTGCATTTGTCATA
TCGTATCTTCATATACCTTTCTTTCGCTATCCTTACTGCTTCTCGACCTT
ATGTGCGTCTGTCTTCCCTTTCTNTATTATTTCTCTGTCTCACCGCTCTN
TGCTCTGTCGCTCCTATCACTAAATTATGTCTCTATCACTGCTACTATCT
GAAGCTGATCTTCGAGATCTCGCTNGGTGTCACTCTTTATCTCATAGNCG
CCTCTGTCTTCTTGTCTCCTTAAGNCTGATTTTCTCGCTCTATTCGTGAC
TACTCTGCTGTCTCTCACATACGTGTTCTTGAATCGTATTCGCGTTCTCG
CTACTGTGATATCCATTGCCGACCTCTACTGCTCNTCTNTATGCTATACT
TCTTAGTCTCTTACTACGTTNGTCTGATATNTTGCTGACGACGTCATGTC
ACGCTCGCAACTCTTCANTTCTATCGTATACGCTGATCATCATTTTCTGT
GAGGCTGATGTACTATACGTAATTACCTGTATACGTCGTCTATCTACTCT
CGTGTCTTCACTCTTTCTACTCC

Seq ID No 91 (>Contig12)
CCCCCCGCCGTCCGCCGGTACGTCGCGGACCGCCGCCCCGAGCAGCTCCC
CGCGCTCGCGCCGGAGGAGCGGGAGGCCGCGGCGCGCCGCCTGTCGGCCC
TCGGCGCGGCGCCGCCGCAGGTCCGGCGCCGCGGGCTGACGCGGGCGCCG
CTCTCGTACGGGCAGAGCCGCATCTACTTCCTCGAGCAGCTCTCGCCCGG
CAAGCCGCTCTTCAACGTCCCGGGCGCGGTCCGGCTCCGGGGCCCGGTCG
ACGTCGCCCGCCTCTCGGCAGCGTTCGGCGAGATCGTGCGGCGCCACGAC
GCCCTCCGCACGTCGATCGCCAACGTCGACGGCGAGCTCCTGCAGATCGC

GCAGCCGCACGCGGGCTTCGCGCTCGACGTGGTGACCTCGACGCCCGAGG
AGGCGGCCGAGCTCGACCGGCGGCTGCGCGCCGAGGCGTGGCGGCCCTTC
GCGATCGGCGCGCCGCCGCTCCTGCGCGCCACGCTGTTCCGCCTCGCGGA
GGACGAGCACGTGCTCCTCGTCACGATGCACCACGTGGTGTCGGACGACT
GGTCGCTCGGCGTGATCCTCCGCGAGCTCCTCGCGCTGTACGCGGGCCGC
TCGCTCCCGCCGCCGCGGCTCCAGGTCAGCGACTTCGCGGCGTGGCAGCG
CGAGATGGTCGAGTCGGGGGCGCTCGACGGCCAGCGCGCGTACTGGCGAG
AGCGCCTCCGGGGGCTGTCCCGGGCGAGCATCTCGGCCGGCGGCGGGGCG
GAGGCGCCGAGCCACGACCCGTCCGGCGCCATCGAGGAGATCGCGCTCTC
GCCGGACAAGGCGGCGGCGCTCGAGGCGCTCGCGCGGCGGGAGGGAGCGA
CCCTGTTCATGGTGCTCCTCGCGCTCCTCGACCTCGTGATCCATGCGCGG
TCCGGCGCACTGGACATCGCCGTGGGGACGCCCATCGCCAACCGGAACCG
CCCGGAGCTCGAGGACGTGGTCGGCCTCTTGACGAACACGCTCGTGATCC
GCGTCGATCTCGCGCGCGCCGGGGCGTTCCGCGACGTGCTCGCGCGGGCG
CGCGTCCAGGCGCTCGACGCCTTCGCGAACCAGGACATCCCGTTCGATGT
CGTCACCCAGGATCTGAAGCAGGAGCGCGACCACGCGCAGCACCCGCTCT
TCCGCGTCTGGCTGGCGCTCCAGAACGCGCCGAAGCCCGCGCTGGAGGTC
CGCGGGCTCCGGGTCGAGCCCCTGCCCCTCCGGCCCGAGCTCGTGCACTT
CGAGGTCGCCCTCCTGCTCTGGCCGGCGGACGACGGATCGGTCGTGGGGC
ACTTCGAGTTCCGGCGCGATCGCGTCGACGAGGGCGCGCGCAAGGAGATC
GCGGCCGCATTCACGCACCTCGTCGACGCGGTGATCGCCCGGCCGGACGC
GCCGGTGTCGACGCTCGTGGAGGGCGCCCGCGCCGAGGCCGCGCGAGCGC
AGGCCGCGCTCGGCGAGGCGTTCGCCAGGGCGGCGACGGCGCGCCTCGGC
CAGCTGCGGCGTCGCTCGGCGGGCGACCGGACGCCCCGCGAGTAGCGGTC
AGCCCTCGGCGGCGGCCAGGCGCACGCGGAACGGCGCAGGGTAGCCGTGG
ACGCGCGGCATGGGGTCGATCGCGCTGGGGACGCCGGCCCGCAGCAGCTG
CTTGATGGCGAGCGAGATGTGCAGGATGGCCACGTACTTGCCGTGGCACG
TATGGATCCCTGACTCCCAGAACAGGTAGTTTCTGCTCGGCAGGCGCCCG
GGCCTGAACTGGTCCGGGGCGTCGATGTGCTCGTGGTCGTGCATCGCCGA
GGCGCTGCAGGCCATCACCAGCGCGCCGGCGGGGACCTTCTCCTCGTGCC

GCGTGCCACGCCCGACCGTGTAGTCGCGCACGCAGAGGCTCGTGACGCCG
GTCGACGGGGACGGAAGCGCAGCGCCTCCAGCACATAGCCGGTGATGGC
GGCGTCGTCCTCGACGTTCACCACGTTGAGCGCGTCGCGCAGGACGCGCG
GGCGCTTCATCAGCTCGACCAGGGCGTTGACGATCGCGCCGCCGCTGAGA
TCCACGCAGCCCATGAGCAGCCCCAGGATCACGTCGCGGATCCCCTCGTC
GCTCTCGTAGGTCTCGGGGACCGACTGCATGACCAGGTAGCGGTCCAGCA
CCGAGGGTTGCTCTGGCGGGGGCGACTTGGCCAGCTGCTTCTTCCGCGCG
GCGACGATCGCGTCGATCATCGGCAGCGCCTCCTGACGAGCGGCCCTCGC
CGCCGCCACGGCCGTCGGGTCGTTGGTCGGGTTGAGGAAGATCTCGTTGA
ACAGCGCGTGGGTCCACGCCACCACCTTCTCGGTCGGGATCTCGCCGACG
CCGAGGTACCGGGCCATCGCGCCGGCCGGCACCCTGAGCGCGTAGTCACC
GGTGAGATCGAACGGCTTGTCGACGCCGACCTTGGCGAGCAGCCGGTTCG
CCTCGTCCACGACGATCTGACGGTAGCGGGGCAGATCGGCGCGCGGGAAC
GCGAGGCGCAGGAGCGACTTCTCGTGCTCGTACTTGGGCGAGTCGTTCAT
CGCCAGGATGTTCTGGCCCACGTTCTCGACCAGCTTGGGCGCGATGTTGT
CGACCGAGAAGACGTCGTTGGCGTTGAGGACCTCGACGACGTCGTTGTAC
CGGGTCACGAGCGTGATGGCCGGGATGGAGAAGATGGGCTTCTCGCGCCG
CAGCTGGCTGAGGAACGGGAGCGGCTCCTCCCTCAGCCACTTGAACACCA
TGCCGGCCTCGATCTGCTTCCGCTTGACCGGATCGTTCTCGTGCGCCAGC
GCGCTGTGGAGGGCCTGCAGGTAATCGAACGGCGGCGCCTTGGCAGCGTC
CGCTCGTCCCTCTTCTTCGATGTGAATGCTCATGGGGAGAATTCCTTTCT
CGCATGCCGATCAGATCGCGACGCTCTGGGGACCATCGACGGGAGCAGG
TACAGGTACGGCTGCTCGCGGGCGCGGTTGCGCTCGGTGATCTCCCGCTC
GATGTGCGCCAGGCGCTCGCGGTAGCGAGCAAACGCCTGCTTCGCCGCCG
GATCTTGCAACAGGCACTTCATCGTCATCAGGCTCTCGCTGGTCAGCGGC
GGGCCCATGCTGAGCGAGCGGCTGAGCACCATCTGGCGGATGCTCTGCGC
GCGCCCCGGGAGGCGCTCCAGCGGCTTGAACTGCCGCTTCTCGCTTCCGT
TCAGGACATCGCCGTAGGAGCGGTAGGTGGCAAACTGAGCGTTCGGGATC
CAGGTGTAATAGTCCGTCTGCCCGAAGTTCACGGCGGCGTGATATGCGGT
CGCCGTGAAGATGATGTTGGTGACGATCGCGATCAGGTCGTCGAGGCTCG

TGAGCTTCTCGAGCTGATCGGCTCGCTCCGGCGGGAGGAGGCTATCCATG
CCGCCGAGCTGGGGGACACGAGCTCGTGGATCCACCGCTGCAGGCTGGC
GTCGCTCGACAGAGACCCCGGCGTCGGGTAGGCGATCTTCAGCACCTGTC
CGACGTACTCCTGGATCGCGTCCCAGTGCAGCAGCGCGTCGTCGCGGTAG
TGATAGCCGACCAGGTCGCGGACGTCGCGCGCCGACAGGTCGCGGGGGAG
CGCGCTCTCGTAGAACCGCCACGGCTTGCCGCCGTACCCTTTGATGCCCT
TGCCGGTGTAGGCGCGCGTCAAGAGCTCGAACGAGCCCATGGTGGCCACC
GAGCTCGTGATGTCGAAGAAGCGCCCTCGCCCGAGGAAGCGCCGGCGAGC
CAGCTCGTTGATGGCCAGGGTGTTGAAGAAATGCGGCCTGAGCAGCTGGT
GGAGCGGATGCGTCGCGGGCAGGTTGCGGTAGGTGCTCACCGCGAACGGC
TCCACGATCAGGTGCGCGTACAGCAGGTGGGTCACCTGGCCCTGGTAGAT
GGCGTCGGCGCTCGCGACGGCGATCTTCGCCGTGAGCCAGTCGTCCGACG
GACCCGAAGGGGTGAAGATCTTGTCGGGATGCGCCCCTTTCCCGGGGCGC
GAGTGCACCAGCCTGATGGCCACGGGCAAGAGCTCACCGGCCGCGGTCTG
GTGCAGCATGCACGTCGGCGCCAGCGGGTACTTGCCCAGCTCTTCCTGCA
CGTCGGTGTCGACGATGTCCTTGAAGATGCGGTAGTCGAGGAAGTAGAGC
TGCCCGCCCTCGCGCACCTCCTCCAGCGTGCGACCGTCGGCGATCGCGAT
CGGCTTGGGCTCGGCGCCGCTCACGAAATCGGCGAGATCGGCCGGGGTCG
CGCGGCGGATGTGCGCCGGGTTGATCCCCACGAGGCGCTGCCGCCCGAAC
TCGGCGTCCTCGGCCCAGCGCGTCGCCACGAGGGGCTTGCGGATGAAGGT
CCACGGCTTGAAGAACTCCTCGAACTGATCGAAGCTCTCCCAGTTGTCGA
TGGACTCGAAGATGGCGCCCAGCCCGAGGTCGGACGTGGCCCTGAGGACG
AACTTCCCCTCGCGATAGCGCTTGTACCCGTACTCGAACAGGTGAAGCGC
CTGCGCGATTTGCAGTCCGGCAGTGTCCTTCCACTTGCCGAGGTTGAGCG
CCAAGATCTTCTTGATCGGAAAGCCTTCCCCCGGCGGTACGGAATCGCTG
CCCGCCGGCAGGTTGGACCCGAAATTGCGCCAGTTCGGTGTCGAGCTGGG
CTCCCTCATGCTCGTGCTCGCTTCTCCGTCTCAGACGGACGGTGGATTGG
GTGGTTCACGTCAAACATCGCTCTCGCGTCGCAGCGGTCCGAGCGCGCGC
CGGAATGGTTCCGTCTCAGTCGCAACAGGACTCAGTACATCCAGCGCCGC
CCCCCGTCCTCGACCTGCCCCCGCAGCCGATCGCGCCGCCCTTCATCGTG

GAATCGACAGGTGCGATTCCACGAAAAGCCGCCGCGCCGAGTTGCACGCG
ACCGATGCTCACGCGTGCATTGTTGAGGCTGCTAGAAAACCGTGGAGCGT
TCACGCATGTCAAGCCATTTTGTTCGGCGCCGCGGCGAGCGGCCGGATGC
CGCGCGCCCCCGCGCCGGGCGTGTTCGCTCCCGACGTACCGCTACCTCGA
CGACGTATGGCTTGAAGGGCAACCGCGCAAGTCGTCCGATTCGTGCTCGT
ATCCTGCTCCTTCCAGCAGGATTTCCCCGCCGCCAGCGGCACAAAGGTGC
CAGGGCGAGCAGAAAGAGCGCTGCCCGCCCCTCCCCGGCCGCGCTCGCCT
CGATCAGCGCGCGCTCGTCGTCGGTGCCTGAGACGTTCGACGGACGTCAG
TTAGTTAGCCTAGCTAACTTCAACACTGATGCGACTGATCGGGCCGACGC
AACCGACGCAACCGACGCAACCGACGCAACCGACGCGACCGACGCAATCG
ACGCAACCGACGTGACGGACGCTGGCGACTCGAAGAAAACCACGGACGCA
CTCCACGTCATCGACGTCATCGACGTCATCGATGCGCTCGATGCAATCCA
TGCACTTGACGCGATCGGTGCGAGCAGGCGACGAGGTCCTCTCGTGAAAC
ACCGAACCGAGTGCCGGTAGCGGGCGCGCCGCAGTGTATGCTAGGCTCGG
CCCTCTTGTCGAGGCCGCGCGCTCGGCGGTCGAGCGTGGGCTCGGGTGCC
GCGGTATCCGGCTGAACCAAGGAGGAGCGAGCCATGCAGGCAGATGACGA
CGCGACGATCTACAAGGTGGTGGTGAACCACGAGGAGCAATACTCCATCT
GGCCGGCGGACCGAGAGAACCCGCTCGGCTGGACGGAGGCCGGCAAGACG
GGCAACAAGGCGGAGTGTCTGGCGTACATCCAGGAGGTCTGGACGGACAT
GCGCCCGCTCAGCCTCCGGAAGAAGATGGCCGAGAGCCCCTGAATCGCGG
CCCGCCCGAGCGCCCGTCGCGAGCGGCCGGGCGGCGGGCTCAGCCGTGTC
ATCGTCGCGCTCGACCGGCCGCGTCCCGCGGATCGCGCGAGCCCGGCGG
GGTCGTGCGCGCCGGCGCTTGTGCCGGGGCCCCCGCTCTCGTACGCCTCC
GTCATGCCGCCCCTCGATCTGCACGTCGCCTTGTTCGGCGCCTCCGGCGC
CGGCAAGACGGTCCTCCTGGCAGCCTTCTACCGGGCGCAGACCCAGCCCT
CGTTCCAGCAGGAGTACGCGTACAAGATCCAGGCGGTCAACAAGGCGCAG
GGCAACCAGCTCCTCGGCCGGTTCTATCGCCTCGAAGAGGGCAGATTCCC
GGACGGCAGCACGCGCTTCGACGAGTACGAGTTCGACTTCTTCCCGAGAG
ATCTGCCCGAGCCGGCGGTCCGCATCCACTGGTACGACTACCCGGACGC
TGGTGGGAGGACGAGCCGGTCGACGCGGACGAGCGGGAGGCGATGCGCCA

GGGCCTCATCCGGCTCGGGATGAGCCAGGTGGGCATCCTCCTCGCGGACG
GCGCGAAGTACCGGGCCGAGGGCACCGGGTACATCCGGTGGCTGTTCGAG
CACTTCGCCGACGAGTGCGACCGGCTGCGCCGGGCCAGCGCCGCCACGGG
CGACGAGGTGAGCTTCCCGCGGGAGTGGATCCTCGCCCTCAGCAAGGCCG
ATCTCTGCCCGCCGGACTACAGCGCGCGGGACTTCGAGCGCGAGGTCTGC
CGGGACGCCGACGATCAGCTGGCGAAGCTCTGCTCGGTGCTCCGCGCCGA
GCACGCGTTCGGCCACCGCTTCATGCTGCTCTCGTCGGTCGCCGCCCCGG
CCGGCGCGCAGGTCGATCCGAGGACCTCGCTCGGCGTGCGCACCCTCGCC
CCCGCGATCCTGGTGAGCACGGTCGAGGGCGCGGTGCGCGAGGCGCAGGC
GGCGAGAAAGGAGAAGTCGGCCGGAGAGACGTTCTTCCAGGGGCTGCGCG
ATCTCGTGCAGTTCGTCGACTCCCTCGACGACTTCCTGCCGAAGCGATAC
CAGATCGTGAGCAAGATCCTGCGGTTCATCTCGATCAAGGACTTCGCGAC
CACCCGGCTCGACCGGCTCAAGAAGATGCGCGAGGACGCGATCCGGAAGG
GCGACACCTTCACGGCGGTCCTGACCGCGATGGTCGCGGCCCTGCGCGAC
GACGAGGGCGCCCGCGCCTACCACCAGAACCAGTGAGGTCGTCATGCCCG
CGCCAGCGCCCCTCGTCGAGACATCGCGCCTCCTCTGGAGGACGCGCGGC
GAGCACTGGGATTACGAGTTCATCTGTGTCCCCGAGATCCCGGCGCTGCC
CGCCTGGCTCTCGACGCTCGAGGCGATGCTCGCCGACGCCGACGCCGGCG
CCGGGGAGCTCCGCTATGGCCTGCTCGAGATCGACGATCGCGGGCAGAGG
GCGCCGCGCGCCTATCCCTACGTGGCCGTGAGGTTCCTCGATCCGGCGCG
GAGGGACTGGACCGGACGGCAGGTCCAGCACTTCGCGGCCTGGTTCCCGC
CGGTCCCGCCCGAGGCGGTCGCGGAGTTGCCAGAAGCGGTCCCCGCCGAC
TGGCACCTTCGCGTGCTCGACGGGCTCGCGGGGACGTACGGCTCCGGCGA
GGTGTTCGGGCTCCCCGAGGCGACGATCCGCGCCTGGAAGCGGAGCCACG
ACGAGAGCCGGGCCGCGCGCGCGATGGCGATCGTCAAGGCGACGCCGCCG
GTTTCGCTGGGCGGCGGCGAGGCGGCGCCGTCGCGGTGGACGCGGGTGCC
GACATTAAAAAAAAAGCCGCCGGAGCCGCCGGCCGCGGCGGGCCTCCTCT
CGGTGGGCGCGGTCCCTAGCGGCCAGGGCCGGCGATTCGGCTGCTTCGCG
ATCGGCGCCATGATGCTCGCCGCCTTCTGTCGACTGATGCTCGCTTGCGG
TGTGCGCCTCCTCGGCGCCTGACGGCTGCGCCGCGCAGGCCATCCGACGG

GGGTCGGCCCGGCCAGCGCCCGCCGGGCGACACCAGGGCATCGGCCCTC
CGCTCGGGGCATCGATTGAGCTCTCCGAGCGGCGGTCCGTCGTCAATCGC
CGCAGAGCTCCCACCGGGCGGAGCAGCTCTGGCCGGTGACCGCATAGGGG
TTCGTCGGGCAGGTCCACCACTCGCCCTGGAAAGGACGCGGGTTGCAGTG
CGGGAGGCACTCCACCCACCCCGACGAGCACGAGTTCCCTACCGAGACGG
TCGGCTGAGCCGCGCAGAACCAGCGTTTTCCTGCGAACCAGCCGGGATTG
CACACATCGGGCGCCCCGCCGACGGGCGGATACACCGTGGCGACCGCCTG
GATGTCCACGGCGTCGAGGGTCTCGCTCCCGAAGCGCACGCCGTCCTGAT
CCGAACACCCACTCCAGTAGGTCATGATGGAGTCGTAGTCGTAGAAACCA
GGGTTCACGACGATGTACCGCCGGCTCGAGGGCCACCCGCTGGCGACGTC
GCTCGCGGGGAGCGGCTCTCGTTGGCTGCAGGCGCTCGGGACCAACGGGT
GATGCCACTCATGCATGAAGCCGATCGCATGACCCATCTCGTGGATCGCG
TACTGCTCCACGCAGTCGAAGCTGTATTCGACCCGGGCTGTCTGCCAGTT
GTACTTGATGCAACGGTTGAAGTCGGCGCCCCAGGGCTTGAACTGGACCG
AGCCGCCCTTGTTGTAAACGCCGATCGAGTCCGATTGGTTGGGCGCGTCG
GGGTGGATCCTGACGCCGACGTAGGTCATGCGAGTGGCCGGCAGGAGCGA
ATCGCAGCTCTCCCAGCCGGTGAAGCGAACCGAGCTCCAGCGTTCCCAGC
TGCCCTGGAGCGCGGTGCGCACGCGCGTGATGACGTCCGCGAGCGAGGGG
TTGGGCGCATGGATCAGCCCGCCCGCGGCGCCGTCGACCCTCTGCTCCGC
CGAGCTCGTGGGGTCGATGCAGACCGGGATCCGGACATGGCCGTCAGCGT
CCTCAGGCCAGCGACTCGCGCTGTCGAAGACGCTCGCCTCGGCGGACCGC
GGCGCGGCGGAGACGGTCAGCGCGGCGCCCAGCGCTGCGAGGAGCAGCGG
ACCGAGCGAAGAGCGAAACCGCACATGTCGTTCAGGGCCCCGCGTCGTGC
GGTGCACCGAGACAATCTCGAGCGGGCTCATGGACGCAAACGCGTTGCGA
TGGCCTTGCAGCATGTTCTTCTCCAATCGACGAGGGTTGTTCTGCTGAAC
GCGGCTCCAGCGTGGAGCTCGACGCGGTTCACCGGCTTCACGCCGGGGCC
GTGGACGAGACCCGAGCACGGGGGAGGTCGCAGCCGCACCGGCTCGCGGC
GCCTCCACCCTGCACCTACGACGAGCCTGCCGCTCGGTTTCGCGGAAAAT
GCCACCCCGCTGCCCAGCGGGCGAAGCGCGGACGAGGCGCTCGTCCCCAC
GGTAGCGCCGGTGCCGCTGCATCCACCGCGCTCCTCCATGGGTCGCTGCC

CGCGGGTCGTCGAGGAGACGGACCCGGGGCGCGGATCCCTGGCTCGGCGT
CGCATAGCTCGTAGGGGCGGCCTTGAGCCGGCGGTACGAGCGGCGCAGTT
CAGCAGCCGACCACGTGGACGCGGCGCGCTCCGAGCTGCGCCGAGGAACC
CTTCAAATATTCAGATGGAATTCACAGGGTGGCTGAGAGACGGGGAGTAA
GATCTCAGAGATCTCCCTGCCTACCCGCATCCCTGTTCAATTTTCCGCCC
ACAACGCGAACGGATGAGGAAATATCAGCCCGCGATCCCGACGGCCGACA
GCATCAAAGGCCGCTCGAATCCAGGGGATTCGAGCGGCCTCGGTCGCGCG
GACCCCCGCCGCGAGCCGCTTTGTCACCACTTCACCACTTCAGAGCTTCG
ATCATCTTCTCACCATAACGCGTGCCCATGATAACAACGGACGCATGATC
GAAGTGGTACTGATCCATGACGTTGGTCCCCTGCTGGCTGACCCAGTACC
CCATAGGCAGCATGTCGGCCGCTTGGTGCACGAGGTTGTTATGACCGCCG
CAGCACCCGCCCGCGGGGAGCTCTCCGAGAATGAAGGGAACGTCGTAGTC
GACCCCCCAGGCTGCTTTCACCTCGTTATAGAGCTGAACGACCTTGCCGG
GCCACGAGCTCTGGCCGTTGTCGGACTCACCCTGGTGGAAGATGATGCCC
GCGAAGCGCGCGTTCTCGGCCGTCTTCGCTTTGGCGATCTTGTTCAAGAT
CATCTGGTGATGCGAGCCACCAGTGATGAACGTGTTGATCGACTCGCCGC
TCTCAGCGGTAGCGACCAACCCGATCGTATCCCCCTCAGGCAGCTTTCCG
AGCAGGGTCTTGCCGAACCAGATGCCCGGGTCGACGGAGGTCGACAGGTT
CCATCCTTTTTCACCAGGGCAATCGCTGAGCGGCGGATTGGCCAAGTTCC
ACTGTCCGGCCGGCTGATTGCATCCGCCGAGGACCTTGAGCCGCGCGTCA
GAATTTTTGTCGCTGTCCTGTTTGTCTGCGACACCAGCCATATTCGACTG
GCCCATGAGCATGAAGATGTGAAACGTCGGACTCGCGCTCGGTGCGCCGC
CGGTGCCTGCCCCGCTGCTAACGGATCCGGTCCCTCCCGTGGCGTCACCT
CCAGTTCCAGCGTTCGTGCTGCCTGTCGCGTCGCCCCCGGTCCCCGCGCT
CGTGCTGCCTGCCGTGGCGCCGCCGGTCCCCGCGCTCGTGCTGCCTGCCG
TGGCGCCGCCGGTCCCCGAGCCGGCCCCTCCGGTGTTGTCGTCCTCACCG
GTCGCGCCGGACTCGCCACAACCGGACGCAGCGATGATGAAGAGGAATGG
GAGGAGCAGGAACCTGGGTGTGCCTCGGGTCGTGCGGTTCATCTCGGTCA
TGATCGTTACCTCGTCGCGCCGGGGCGCGATCTGAAGAGCATGGCGGAAT
CGGTAGGCCGGCGTCGCGATGCCGGCGCGGCGAACCTCGCCCGCAAAGAG

CTCAGCGCCGGGGCCTACCTTATCGCATCTTGGGCGCTTGGCGTCCAGGA
TTCGGCCTTAGACAGCACAAGCAGAAGACCTTTGACACTGGATTTTTTCA
TCATCGGCGGCGCTCGTTCTTCGCTGCGCCTCAAGCGCCGACCGTTCGTT
TCGAAGCGAAGCGGTTTCACGATCCGGGATCGCGGAAATTTGAAACGGAC
GCGTCGCGCGGGCAACGCAGGGGACTCATCACGAGGCAACCGCGCTGCGT
CGCGAAATTGGCCAGCCTCTCGGAGTCCCTAGTTCCGTGCGTCAGACGCG
TCACCCACCATGTCGAGCTCGGCGCGGCGCTCTACGTGCTTGAAAGACCT
CCGCGAGCGCCGCGCTCTGCGCTCCGCGAGCGCACCAGGCTCCCCGTGGA
TTCAGGGCAAGGCGGTCGTGATCACGTCCTTCGTGGTCGCCGTGCTCTGG
CCGGGGCGGTCGACGCAGGACGGGTAGTACTGGCCGGCCCCGTTGAGCTT
CTGGCAGACCCGGTCGCACGACCCGGCGATGCGGATCAGCCCGCACTCCA
CGATCTGACCCCCGGAGGTCACGTGGCCCGCGGCGCAGTCGCGCTGGTAG
GCGCGCGAGTTGTCGACCGTCGCGCTGTTGTAGCAAGCGTTGATATAGGG
TTGCGCGGCGAACAGGTTGCCCCAGAAGGCCCCCTCGACGTCCGGATAAT
CGATGAGCTCCTGGCTGGAGGAGAGCGTCTTCAGCGGATCCCGCAGCGAG
CGGGCGGAGAGGAGCACCGGTACTTGATAGTAGTTCACGCGCGCCGCCAC
GCAGCTGGACACGATGCGCTGCCCTGCGTCGTCGAGCGGCCCGCTCGCCC
ACGCGGGCGCGACGCCGAGCAGCCCGGGGTAGCGCTCGTCGTGCCTCTTG
CCGTTCGAGTCCGTCCACGAAAAATCGAAGGAGGCCGTGCTGCTCAGGGC
GCAGCTCGCCGCGTAACGCAAGAAATCGCGCGCCAGCGCGCCGCTCGGCC
CGGGATCCTGGATCGCGGCGAGGTTCCGCGCGCTGAGGCCGCTCAGGTTC
AGGGCGTTCAGGTTCAGGGCGTTGAGGTTCAAGGCGTTGAGGTTCAGGGC
GTTCGTGCTGAGCGCGTTGCCGCCCACGAGGGCCCCCTGGGATTCCCCCA
CAGGCTCGCCCCACGCATCGGCGTCCACCACCTCGGCGGCGCAGCCCGAC
AGCACCCCTGCCCAACCAAGCACGATGAATGTCCGCTCGAGAGACATGGA
TTCCCCCGTGTTCCTGGCGCATGACCCGACGGCGCCCTGCGCGCGGCGCG
CGCGGCTCCCATCGATTCGCTGGATGGGTTCAATATTCTACTTTTTCCC
GCGCTCTCGCGCCGGTGAAAGTCGCTTCAGCGGCGGCGAGGTCGATGTCA
GGAGCGTCCGACTCCGTCGCTCTCGTCAGCTCCGCGTACCAGCGACGGAG
TCGCCCGCCCATGACGGTCGGAATGGTAGAGGCGGCCGCGAGGGCGCGCT

CGAGCTGCGCCCGGGCGTCGGCGCGCCGGCCGCGGCGCAGGGCGGCGAGC
GCCCGCGCCTCGAGCACCTCGATCCGCTCCTGCCCGACCGAGCAGCGCGC
GGAGCGCTCCTCGAGCGCGGCCCACGCGGCGCGGTCGTCGTCGCGGGTCG
CGAGCTCGATCATCGCGCAGAGCACGTCCTCCGAGGGCTTCAGCGCCTCG
CAGCCGGCGTCGTCTCGCGCCGCGCGGAGCCGCAGCGCGATCCGGCGAGC
GCCCGCCTCGTCGCCCTGGTAGAGGCGCAGGCGCGCGATCAGGAGCGTTA
CGACGACCGGCGCGTGGCGATCGCCGCAGCGCGGCGCTATCGCCTGGACG
GCGCGCGCATGGGGCCTCGCGGCCGCGAGATCGTCCATCAGGTACAGGTA
CTCGGCGAGGTTGTAGCGGCCGACGAGCTCGAACGCGGGCTGGCCGAGCT
CGCGCCCGAGCGCGATGGTGCGCTCGAAATCGGCGATCATCCCGGCGCGA
TCGCCCTGGAGCGCCCGCGCGAGCCCGCGGTTGTTGAGCGCGGCGCCGAG
GTGCATGAGATCGCTGCGCTCCTCGCAGCTGAGGATCACCGCGTCGAGGT
CTCGCGCCGCCTCCTCGACGCGGCCGAGGCTGGCCAGGATGAAGCCGAGC
AGCAGCAGGGCGATGATGTGCGTCTCGTGGCCCTCGTCCCCGAGCCGCGC
CGCCTGCGCCGCGGCGCGCGTCAGCACCGCGGCGGCCTCGTCCTCGCGGT
CGGCGCGGTGGAGCGAGCGGCCCACGCCGAGGAGCAGGCGGGCGCCGAGC
AGGGGCGAGGCCACCCGGCCGGCGAGGCGCTCGGCGGCCGCGACCCGCTC
GCGCGCGGCCCGGTACTCGCCCGTCCAGTCGAGGATCATGGCCTCGTCGA
GGAGGAGCTCGATCTCGGCCCCCGCCTCCGACGCCGCCGCCGCCGCCTCG
CGCGCCGCGGCGAGGTCGGCGAGGGCCTCGGTGTGGCGCCCGAGCCGGAA
GCGAGCGAGGCCCCGCGCTCGGCGCTCCTCGGGGAGCAGCGCGCCGAGCA
GCGCCTCGACGCGCCCGTAGCAGCCCTCGGCGTCGAGGTAGGCCCGGCGC
GCGGCCGCGAGCTCGGCGCCGCGGGCGAGGAGCGACGCCGCGCGGGCGGT
CAGGCCGCCGCGCTCGCAGTGCGCCGCGAGCACCAGCGGATCGGCCTCGC
CCGCGGCCTCGAGCCAGTCGGCGGCGAGGCGGTGGCCGAGCGCGCGATCG
GCAGCTCGCCGCGTAACGCAAGAAATCGCGCGCCAGCGCGCCGCTCGGCC
CGGGATCCTGGATCGCGGCGAGGTTCCGCGCGCTGAGGCCGCTCAGGTTC
AGGGCGTTCAGGTTCAGGGCGTTGAGGTTCAAGGCGTTGAGGTTCAGGGC
GTTCGTGCTGAGCGCGTTGCCGCCCACGAGGGCCCCCTGGATTCCCCCA
CAGGCTCGCCCCACGCATCGGCGTCCACCACCTCGGCGGCGCAGCCCGAC

```
AGCACCCCTGCCCAACCAAGCACGATGAATGTCCGCTCGAGAGACATGGA
TTCCCCCGTGTTCCTGGCGCATGACCCGACGGCGCCCTGCGCGCGGCGCG
CGCGGGCTCCCATCGATTCGCTGGATGGGTTCAATATTCTACTTTTTCCC
GCGCTCTCGCGCCGGTGAAAGTCGCTTCAGCGGCGGCGAGGTCGATGTCA
GGAGCGTCCGACTCCGTCGCTCTCGTCAGCTCCGCGTACCAGCGACGGAG
TCGCCCGCCCATGACGGTCGGAATGGTAGAGGCGGCCGCGAGGGCGCGCT
CGAGCTGCGCCCGGGCGTCGGCGCGCCGGCCGCGGCGCAGGGCGGCGAGC
GCCCGCGCCTCGAGCACCTCGATCCGCTCCTGCCCGACCGAGCAGCGCGC
GGAGCGCTCCTCGAGCGCGGCCCACGCGGCGCGGTCGTCGTCGCGGGTCG
CGAGCTCGATCATCGCGCAGAGCACGTCCTCCGAGGGCTTCAGCGCCTCG
CAGCCGGCGTCGTCTCGCGCCGCGCGGAGCCGCAGCGCGATCCGGCGAGC
GCCCGCCTCGTCGCCCTGGTAGAGGCGCAGGCGCGCGATCAGGAGCGTTA
CGACGACCGGCGCGTGGCGATCGCCGCAGCGCGGCGCTATCGCCTGGACG
GCGCGCGCATGGGGCCTCGCGGCCGCGAGATCGTCCATCAGGTACAGGTA
CTCGGCGAGGTTGTAGCGGCCGACGAGCTCGAACGCGGGCTGGCCGAGCT
CGCGCCCGAGCGCGATGGTGCGCTCGAAATCGGCGATCATCCCGGCGCGA
TCGCCCTGGAGCGCCCGCGCGAGCCCGCGGTTGTTGAGCGCGGCGCCGAG
GTGCATGAGATCGCTGCGCTCCTCGCAGCTGAGGATCACCGCGTCGAGGT
CTCGCGCCGCCTCCTCGACGCGGCCGAGGCTGGCCAGGATGAAGCCGAGC
AGCAGCAGGGCGATGATGTGCGTCTCGTGGCCCTCGTCCCCGAGCCGCGC
CGCCTGCGCCGCGGCGCGCGTCAGCACCGCGGCGGCCTCGTCCTCGCGGT
CGGCGCGGTGGAGCGAGCGGCCCACGCCGAGGAGCAGGCGGGCGCCGAGC
AGGGGCGAGGCCACCCGGCCGGCGAGGCGCTCGGCGGCCGCGACCCGCTC
GCGCGCGGCCCGGTACTCGCCCGTCCAGTCGAGGATCATGGCCTCGTCGA
GGAGGAGCTCGATCTCGGCCCCCGCCTCCGACGCCGCCGCCGCCGCCTCG
CGCGCCGCGGCGAGGTCGGCGAGGGCCTCGGTGTGGCGCCCGAGCCGGAA
GCGAGCGAGGCCCCGCGCTCGGCGCTCCTCGGGGAGCAGCGCGCCGAGCA
GCGCCTCGACGCGCCCGTAGCAGCCCTCGGCGTCGAGGTAGGCCCGGCGC
GCGGCCGCGAGCTCGGCGCCGCGGGCGAGGAGCGACGCCGCGCGGGCGGT
CAGGCCGCCGCGCTCGCAGTGCGCCGCGAGCACCAGCGGATCGGCCTCGC
```

CCGCGGCCTCGAGCCAGTCGGCGGCGAGGCGGTGGCCGAGCGCGCGATCG
TCCTTGGTGAGCTGCGCGTAAGCGCCCTCGCGCAGGAGCGCCTGGCGGAA
GGAGTACTCCTCCTCGCCGGGGAAGCGGCCCTCGCGGTGGCGGACGCAGA
GCTCCCCGGCGACGAGCGCGGAGAGGTGCTCCGCGAGCGGAGCGGCCTCG
TCGCCCCCGAGCAGGTGCGCGACGGCGCCTCGCCAGAACACCTCGCCGAG
CACGCTGGCGGCCCGCAGGATCCGGCGCGCGGGGGCGCGAGCGCCTCCA
GCCGGACCTGCACCATCGCCACCACCGTCTCGGGCAGCGCGTCGCCGCGG
CCCTCCGCCGTCGCGCGGATCAGCTCCTCGAGGAAGAACGGCTGGCCCTC
GGACTGGGTGACCAGACGATCGATGAGGGCCCCGTCGGCCGCGTCGCCCA
GCGCCTCCCGCGCGAGCTGCGCGCACGCCCTCGGCGGGAGCTGCCTGAGC
CAGAGCTCCTGCCGCCCGCGCTCGGCCCAGAGATCGGGGTACGCTTGC,

or their complementary strands,

(b) DNA-sequences which hybridise under stringent conditions to regions of DNA-sequences according to (a) encoding proteins or to fragments of said DNA-sequences,

(c) DNA-sequences which hybridise to the DNA-sequences according to (a) and (b) because of a degeneration of the genetic code,

(d) allele variations and mutants resulting by substitution, insertion or deletion of nucleotides or inversion of nucleotide segments of DNA-sequences according to (a) to (c), wherein the variations and mutants offer isofunctional expression products.

13. Peptide encoded by a DNA sequence according to claim 12 selected from the group consisting of

Seq ID No 92

>Contig11_002 591 amino acids MW=63639 D pI=5.80 numambig=0
MLDVWSTSDQVACRLHCAGAGPSASLELRYDASAGARRDAERLAERLAALLEDLSRHPER
PVAQGEYVGPGERAEIEAWSRGPAMELPSACALHRWFEERAEQHPDVVAVRSEGKSLTYG
ELERRANRLASCLRRRGVGLDTIVGVCVPRSEDMVVATLAVLKVGGAYLPLDHEYPGERL
AFMMRDARARLLVTHDAIADELPTGGWTTLLLDAEAAEIAACSDARPAVSPPPDSGAYVI
YTSGSTGTPKGSLISHRAIVNQMQWIQRYWALTADDRVLLKAAFGFDVSVWEIFWPLSFG
ARIVVARAGGHRDPEYLRRLVRDEGATTAYFVSSMLAAFLGGPEQPFPASLRKVLVGGEA
VPLDLVRRFYAKHDGDLINMYGPSEAAIAVTGCVLPSDPRVTWVPLGAPVANAEVFVLDG
AMRRPAIGALGDLYIAGAPLARGYVGQPGLTAERFLPDPCARAAGGRMYRTGDVARFLPD
GMLEFQGRSDHQIKLRGHRIELGDVEAQIRRVPGVGQAAVVLREDAPGDARLVAYVVLDG
DAAGDAPDVRAGLKASLSAYMIPSSVVRLYALPMCSERLAFTGSSYAGCLL*

Seq ID No 93

>Contig11_007 361 amino acids MW=38862 D pI=10.42 numambig=0
MSDHEMTGFSLSPQQRAIRALDREAGAPGCRTLAVVAVTGPCDEGRLSAAALALAERHEI
LRTRLVEAGRARPRRWSASRASRGRQQDDWVGCSEAEQGERMSRLVARLSEDRGADDGLR
VGLVRVGPEERRLVLAAPAWCVDEESIAPLVRELCASTAGAGAPPEQQYADVAEWLNGML
ESEDAGDGRRFWAERRSHFGPPLHLAFSRGGAGAGAGSGRARVDLGQGGMAQVERWSSSW
QVPQRIVLLALWASLLWRMSGGNEPEVTVAVRFDGRSLDALAGAVGPFARFLPVRIEISA
SDTLADVARRLALAEAEAAAHQDAAPGVSHRMSWGLLRRGGRAGAVARRRAGPRARRLEH
V*

Seq ID No 94

>Contig11_012 882 amino acids MW=95015 D pI=12.69 numambig=0
MARALYAQEAAFRSAFNACAAAARARGLDLGEVVFREGEGDGRTLLGTALAQPALFAVEL
ALARLWMSWGIEPAAMIGHSLGELVAACVAGVFTLEDAMSLVIDRGRFMQAAPAGSMLAV
GLPAADVEGLLEAGLEIAAVNSPKLTVVAGPASAIRDLAARLEAREVFARPLQTSHAFHC
ALIDGAVAPFLESVRRARLSPPEIPVVSNVTGALLTDAEATDPAYWARHLRQPVRFSDGV

EALFASGHALFLEVGPGRGLTTLVRQTLAGRGGAAIASLGSTHAASEPASLAEALGQLWE
AGHAVDWTARPRGRPPARSSACRRTRSSGRGTGSRRAAAPPPPSRRRPTRPSRPRRTPSR
RRRSPARTPGPRSPRRRPSPSARSPAIWERPPRRPRDRRPRRLLRPARRLARRDPGDVPD
PRAARRRAPAAGALPGADGRGARGPGRRRAPRRGAAARVPPDRADPPRRPAPAVVRAAPA
VVRRSARAGQPGVQHPVRGARDGPARRRRAPPQPVRDRAAPRGAADDVQRQGRRAVPGRR
ARGPGAVPDVRPRAPRGRGPRRGGLGARPRGEPRAVRPEPRAAPPRAGHPEAPRRARHRA
RRAPRRVRRLVGRRLRRRARRALRRLRAGPAFPAAGAPGAVRRLRRSPAGVAVRRGPGGR
APVLDDEALGRAAPGPRAGRPRAGGAKDVARGAAEPRRRGRAHPPDQGVLRAGGDLPLHG
PARRVQARPAPAHRARGPRRGHRRREPEPRRDRADDRLLRQPARAPDRLRGRPDVRRAGP
PRARRRARGVRAPGPAVRPARRGPAAEGGRRARAPLRREVRHAQRARPPDEARGARARGA
GGRGDHDGVRLRPHGRRGRRLVPLRRRAQQRAVPGGDGRQFP*

Seq ID No 95

>Contig11_021 1213 amino acids MW=131017 D pI=12.40 numambig=0
MRGRRRRAAPHLRGARPARRGDRRGGLALRSDRERARRRGHGEGLGAGGRRARRAARRRR
LPAPGPSPPRGAPQAPPGARRGPPRPHAVGGRRDDRLARGHRAPRRRRGRAVARAARRAP
AARRKHGRPGLRDLHVRVHGAAQGRDDRPPRRGEHGPRHQPPLRRRPGGPGARALVAELR
PVGLRRVRDARRRRRRRDPRPYPRLGSGALARARGARAGDRVELGPGADGDAHGRVPRRG
RPGAVVAPPRHDERRLDPAEAPRSHPRGLPRAPRREPRRRDRGVDLVDRPPDRGRRPGVA
QHPLRPPAREPAHLRARRGAGAVSDRGPRRDPHRRDRRRARILARRGADPGAVPEAPHDR
RAAVQDRRPGPLLRRRDHRAARAHRPSGEDPRLPHRARRDRGRPRAAPLGRAGGRGGEDR
SVGREAPGRVRRRRRRRWRRAPRLRPEEAAGVHDPRGGRRPPGAAAERERQGGPRRPAGS
RRGRAARRGRRPADRDRAAHRLGARGGLAGRGRRRDRQPLRARLHLAAPRARAAPARRAH
RGARARRGRRGAGRVADRPVPVPDHRATGAAARRGDGQGRAGRRRGAARRGAPRRATPPG
ARIAMSEPIETEDGGSDIAIVGMAGRFPGAPSVDALWENVRRGVESIARFPESEREEPPV
GASAAPGAPVVCAGGLLDDIDRFDASYFGYSPREAQLMDPQQRLFLECAVAALEDAGCDP
ARFPGAIGVFGGCGSNTYLLQLLSHPDLAATVDPHALMLASEKDYLATRVSYKLDLHGPS
VVVQTACSTSLVAVHMACESLLGGQCDLALAGGVSIGIPQKRGYPYVPGSICSPDGRCRP
FDARAEGTVGGSGVGIVALKRLADALRDRNTVHAVIRGSAVNNDGGRKVGFMAPSVDGQA
AAISEAQSVAGVDPGSIGYVEAHGTATAIGDPIEVEALTQAFRRKTPRKAYCALGSIKAN

IGHLDAAAGVAGLIKAAHVVRSGEIPPCVHFEAPNPKLDLAASPFFVPREAAPWPRELRP
RRAGVSSFGIGGTNAHVVLEEPPPLPPRAPAPERDHVLTLSARTPEALSTACAQLAAHLE
ATDVPLDDVAFTLQTGRAEHPYRRAVVARTRAEAIQGLAREGASALARPDEPRPSSRSRA
RARRPSGWPARSTRRRRRSGAPSTRARRRRGRAASISARSSSARARATGARCSAPRWRSP
RSSPSSSRSPGSG*

Seq ID No 96

>Contig11_026 3079 amino acids MW=332984 D pI=5.97 numambig=0
MLTVVDHHVVVEYWSFALIVRELGELYSALRAGRPPQLPPPSSFFAAGVSCPSPREAAGG
AEYWRKALDGATTAIDLPRDRARHDAGARRGRAHAITLPKPLTGALARLARERGTTLFSV
LLSALTVLLHRASGQSDLVVGVPSAGRHDDESARAFGYFVQMLPVRVALRGAASFDALVA
RVRDAFLDALAHGDSALRHLLARAQGEAQRDALFDVAFAFQSTPPSLDARLSALAIGVGD
VRIAQGELELTTLADEQAAAEFDLALFAAELDAGIALRFEYDQQLFDPATIERMARHFVV
LLESAVEHPGRPLSELRMLSDAERALLLDDWSGAAAARQAASAPAPACVHALFEAHAARQ
PDATALEFGHQRFTYAQLSTWSTELALWLRDRGVGPGSVVGVCIERSPRMVAAQLAVLKA
GAAYASLDPANPPARLAEMLADCRASLALTSSQASHKLTAAPCPVHLVQDGACAPSTHIP
LVSRPDDLAYVLFTSGSTGTPKGVCVRHASLSRLVSFLHLRLDLSPSDRWTQVASSGFDA
SVYEIWTPLACGAALLLADDDALRSPTALVSWLVAQRATLSFMPTPLAEACFEQDWTGSA
LRAMTVGGDKLHPLRRPLPFRLFNMYGPTEATVITTVAEIADLGAEPPLGRPVDSALVYV
LDPHMQPVPPGALGELYIGGACLAQGYTRTDLTAERFLPDPFGQPGARLYRTGDLVRWRP
DGQLAFAGRRDEQVKLRGRRVELGEVESVLRRLPGVREGIVVLHGQGSAAHLIAHVVPDA
HPPSERDLREGMARLVPDALVPAHFVLLPALPMSLSGKVDKKLLPAPPAAHADYEPPSGE
LELELAHIWQSVLHLDRVGRHDSFFDLGGHSLLAMQVLGRIESSLGIRTTLRTLFEHPTL
AQLASHLSSGAASTSAAAATALERGLTRPDGPSSPRVATPEEPFALTEGQRAMWLECQKS
ADGALYNLGRTVRLGAGVDVAALRRAFEGLVERHEALRTTFLTRDGHPLQQVHRHVALEW
AEEPAMALDEREIVARADEVRRRAFDLERGPLLRVHVWRRGEGQPPLLTVVVHHLVVDYW
SFALLVRELGELYSALRAGRPPQLPPPSSFFAAGVSCPSPREAAGGAEYWRKALDGATTA
IDLPRDRARHDASPRRGRAHAITLPKPLTGALARLARERGTTLFSVLLSALTVLLHRASG
QNDLVVGVPSAGRNDDESTRAFGYFVQMLPVRVALRGAASFDALVARVRDAFLDGLAHGD
SALQHLLAEPRGAARRGGALFDVAFAFQGALPSLDPRLAALTTGAEDVRIAQGELELTTL

ADEQAAAEFDLALFAAELDSGIALRFEYDQQLFDPATIERMARHFVLLLESAVEHPGRPL
SELRMLSDAERALLLDDWSGAAAARQAASAPAPACVHALFEAHAARQPDATALEFGHQRF
TYAELSTWSTELALWLRDRGVGPGSVVGVCIERSPRMVAAQLAVLKAGAAYASLDPANPP
ARLAEMLADCRAALVLTSSQASHKLTAAPCPVHLVQDGACAPSTHIPLVSRPDDLAYVLF
TSGSSGTPKGVCVRHASLSRLVSFFQHLLALSPRDRWTQLASSGFDASVYEIWTPLACGA
ALLLADDDALRSPTALVSWLVAQRATLSFMPTPLAEACFEQDWTGIALRAMTVGGDKLHP
LRRPLPFRLFNMYGPTEATVITTVAEVADLGDEPPLGRPIDSALVYVLDPHMQPVPPGVL
GELYIGGACLAQGYTRTDLTAERFLPDPFGQPGARLYRTGDLVRWRPDGQLAFAGRRDEQ
VKLRGRRVELGEVESALRRLPAVREGVVVLHGQGSAARLIAYVVPGADPPSERDLREGMA
RLVPDALVPAHFVLLPALPMSLSGKVDKKLLPAPPAAHADYEPPSGELERELAHIWQSVL
HLDRVGRHDSFFDLGGHSLLAMQVLGRIESSLGIRTTLRTLFEHPTLHQLADRLSSGAAS
TTAAAATVPASEIAPSLGRAPADEPYPLSYEQERLWVLEQLLPGGTAYNVVQAVRLRNLV
DVDALSSALAALVRRHWSLRTVFVASPTPAQKICEPEAAPAEVVDLRGTPPDEAEAAARA
WASREQATGFDLARGPVFRARLFRLDHDVCVLVLSTHHIVTDAWSFQPLVRDLAELYRRA
RGGGPADMPELPLQYVDFAVWQRRHLAGKRLADKLAHWTATLRGLPVLELQTDRPRPPVQ
TFRGAERVLPLDARLVAQLDELARSRGATRFMVLLAALGVLLRRSSGQDDLAIGTAVANR
PRPELEPLVGFFVNTIVMRLDLGGDPTFEELLSRARKVALEAFEHQDAPFEKVVEAVNPR
RDLSRSPLFQVMLVVQNAPTEALELGEVRIEPLDLPVEATRFDLRFSVEPRGGRDVISLQ
YNVDLFDAATIDRMLATMQSVLSRATQDPAQRVRALSVAPEDRERALVAWNDTAVATPDH
LRLEEPFFERAVEQPDACAVVDAERRLTYGELARRAEAIAAAASRSGATANALVAVVMEK
GWEQVAAVLGVLRAGAAYLPLDPRLPEERLRHLLEHAEVRLVLTQSAVDGTIAWPAGIER
LAVDADERWREQPVARRPPGGSTDDLAYVIYTSGSTGLPKGVMIDHRGAVNTVLDINRRF
DVGPEDRVLALSSLSFDLSVYDVFGTLAAGGAVVIPDRTRASDPGHWRELVERERVTVWN
SVPALMEMLMDASPGAGDPALSSLRLVMMSGDWIPLKLPDRIRAACRAPRVVSLGGATEA
SIWSIAHPIADVDPAWRSIPYGRPLANQHTYVLDEGLEPCPIGVPGEIHIGGIGVALGYW
RDEARTRERFLKHPTTGERLYRTGDLGRYFADGTIELLGRTDHQVKIRGFRIELGEIEAA
LAQHPSVEQAVVAAKTDPSGEKRLVAYVVGADGDGAALRDFVRKKLPEYMIPAEVVVLPA
LPLSANGKVDRAALPDPAAVAPRAAAVAPRTATERLIASVLAEVLQVEAVGVTDNLFELG
FTSLLLVRAQRLLAERIAARAPDEGAAAQAVSLTDLFQYPTIEQLAQRLDAATVKAEPAD
VGAQRAEARRDARRRRGRG*

Seq ID No 97

>Contig11_011 544 amino acids MW=60164 D pI=9.10 numambig=0
MMSRIRAQLGVELPLRALFQGPTVAALAAQVDAARRGEARRREFPPIARIPRDGPLPLSF
AQHRLWFVDQLEPGSPAYNIPFVVRATGRLDVDALRRSLFEIARRHEALRTTFSARDGVP
FPVVAPEARVPFRMSDLEHLAGEALDAAVSALVLEESLAPFDLSRGPLLRVRVIRKRHDE
HVIALVVHHVVFDVWSVGVFVGELAALYGGFAQGQPSRLPELPAQYVDFAAAQRAWLSGE
VLEGELRYWTTKLSGALRRARVPVDHEPAGRRTWRGARRSLDAGAELTRQIKAFCEREAI
SPFMALLAAYKLVLHQRTGLEDLVVGTDVANRNRVETEPMIGFFVNQLVLRTDCGGDPTF
GALVRRVRDVALEAFEHQDLPFDRLVEALRPKGAVGHVPLFDAKFVMRNVHVPPMKLEGL
ELEALEGEATTTAFDFVLTVAEAGGSFRFGVEHSSELYRAATVDNFLSDYRQILATATAR
PDTPVSELRGELERAAAARRELERKAARGAALDKLTSARRRAVTLPRPGAPGEAKTSPKD
DLDE*

Seq ID No 98

>Contig12_001 514 amino acids MW=56145 D pI=8.82 numambig=0
PPAVRRYVADRRPEQLPALAPEEREAAARRLSALGAAPPQVRRRGLTRAPLSYGQSRIYF
LEQLSPGKPLFNVPGAVRLRGPVDVARLSAAFGEIVRRHDALRTSIANVDGELLQIAQPH
AGFALDVVTSTPEEAAELDRRLRAEAWRPFAIGAPPLLRATLFRLAEDEHVLLVTMHHVV
SDDWSLGVILRELLALYAGRSLPPPRLQVSDFAAWQREMVESGALDGQRAYWRERLRGLS
RASISAGGGAEAPSHDPSGAIEEIALSPDKAAALEALARREGATLFMVLLALLDLVIHAR
SGALDIAVGTPIANRNRPELEDVVGLLTNTLVIRVDLARAGAFRDVLARARVQALDAFAN
QDIPFDVVTQDLKQERDHAQHPLFRVWLALQNAPKPALEVRGLRVEPLPLRPELVHFEVA
LLLWPADDGSVVGHFEFRRDRVDEGARKEIAAAFTHLVDAVIARPDAPVSTLVEGARAEA
ARAQAALGEAFARAATARLGQLRRRSAGDRTPRE*

Seq ID No 99

>Contig12_009 582 amino acids MW=65555 D pI=8.72 numambig=0
MREPSSTPNWRNFGSNLPAGSDSVPPGEGFPIKKILALNLGKWKDTAGLQIAQALHLFEY
GYKRYREGKFVLRATSDLGLGAIFESIDNWESFDQFEEFFKPWTFIRKPLVATRWAEDAE

FGRQRLVGINPAHIRRATPADLADFVSGAEPKPIAIADGRTLEEVREGGQLYFLDYRIFK
DIVDTDVQEELGKYPLAPTCMLHQTAAGELLPVAIRLVHSRPGKGAHPDKIFTPSGPSDD
WLTAKIAVASADAIYQGQVTHLLYAHLIVEPFAVSTYRNLPATHPLHQLLRPHFFNTLAI
NELARRRFLGRGRFFDITSSVATMGSFELLTRAYTGKGIKGYGGKPWRFYESALPRDLSA
RDVRDLVGYHYRDDALLHWDAIQEYVGQVLKIAYPTPGSLSSDASLQRWIHELVSPQLGG
MDSLLPPERADQLEKLTSLDDLIAIVTNIIFTATAYHAAVNFGQTDYYTWIPNAQFATYR
SYGDVLNGSEKRQFKPLERLPGRAQSIRQMVLSRSLSMGPPLTSESLMTMKCLLQDPAAK
QAFARYRERLAHIEREITERNRAREQPYLYLLPSMVPQSVAI*

SEQ ID No 100 (>ORF1)

VSQRTSCYLRGGGVCSMNDAFLALERNERNRPSTVIDLLRQRAEAEPARPIYCFLESGDVEAG
ATWVTLREIDERARTVAALLQASGVAPGARALLLYPPGIEYITAFFGCLYAGVRTVPAYPPDL
GRLERTLPRVASIVADARAEAALTSSAVAGIVASLPASAAAAALQRLRWIATDGPSPGPIEGP
GAALRPESVAFLQYTSGSTGEPKGVMLTHGNLLHNSRLIAHGFDLTSPDPVGVIWLPPYHDMG
LIGGILQALYRRIRVALMSPLSFLQRPMRWLRAVSALGASVSGGPNFAYDLCVRKSSEEERAA
LDLRSWEVAFTGAEPVRADTLDRFARAFAVSGFRREAFYPCYGLAEATLIVSGGARAEAPVLA
RLAPEEVELGRAVASAAEGARVFVGSGRALDPRAVAIVDPAGNELGPGEIGEIWVSGPSVAVG
YWGRPEETEATFGATLAGSAAPRYLRTGDLGFLRGGELFVVGRSKDLIILRGRNHFPQDIEKT
VESSHRAVRPGCSAAFSVEHEGEERLAVVCEVDPRVAADPREIVAAREAVTAEHQLVAHAVAL
IAPGALPKTSSGKVRRRECRRAFLEDALGERHVAFAPELLDDASPPDDAPPETEEPSGRSLLD
ALRSTLARALRLDAGQIDDALPISRFGLDSLAAVELQHAFQVRTGRAIPLTSILRGGSLRLTR
EITRLDGPSSPRVATPGGAVCADRWGTGRFGSSAISRPMERFTTWAGRSGSVPAFKRVDLRRA
F

SEQ ID No 101 (>ORF2)

VYSSAYVLFAVCAGTTRVASAPETAGFPLECVGDDGTVLGPDSFVVGYTQVYVFKKERLNTNP
PIDGFTLKLDGNEVAPGEDGLPVVKRCVRSEEQAQGCGRTEPAEDECTTYEIEAVVPEKAAEV
DEEAAGLGGPPAREAIWVDYYTDGGEFDGARRLVSDTTGASRGGNGTTWTPPSEPGRVSLWAV
VHDTRGGASVTRREVQVE

SEQ ID No 102 (>ORF3)

VVGTVLSAGTGEPLPDIAVTLVRPDGGREEAKTDQGGKFRFKNLPPGKYRVEVAAAGFEPFAA
EEEIAAGEAIEVRYRISLAAPQDGKAPGIEVTVQGERPPREVTRRTIERREIDRIPGTGGDAL
RSLQSLPGVARSGFGLLIVRGSAPQDTLTFVDRTPVPIIYHFGGLSSVVPTEMLEKIDFYPGN
FSAVYGRAMGGIVDVGLRSPKQDGKYHGVVQLDLIDGRVLLEGPVPFLKDWTFIAAGRRSWVD
AWLGPVLKEAGSSVTQAPVYYDYQFVLEGRPSASERVRASFYGSDDAFKITLDKPPEDEPALT
GDFGLHTAFQRFQLSYENRIGSRDRLLWSMALGRDIADFEISPLAFNVVSTSLDLRLELSHRF
ARYLTMNVGTDLSGGVATVNIRAPSQQPAGHPSNQPFSTYPFQDRSFDGAYSRPAAYAELEVV
PSPRARIVPGVRVDYALDTQTLDVSPRVNARYDIRSGFPRTTAKGGVGLYYQAPQFAESIEPF
GNAELKSNRAVHYGLGVEQEITPQIEVTLDGFYKQLDRLVVFSPEKDDYADGTGYAVGGELLL
KYKPDERFFGWAAYTLSRSVRKDGPDEEEHLTQFDQTHVLTVLGSLRLGRGWELARFRLVSGN
LQTPYVCDPEEKGCNPNRVNAIYHASSARYSPIPLGGDYSERMPLFHQLDIRADKTWKFKRWQ
LGLYLDIQNVYNYMAAEGISYNFNYTKREYVTGLPFLPTLGLRGDF

SEQ ID No 103 (>ORF4)

VIAVDNNPEAVDAVKDKTSAAFVGDATVHKVLEGIGAQYVETAIVTFGEHFEPSVLCVASLVR
MGVRIIARAATDRQADILRAVGATRVIQLETEMGRRVGADITMPLAQDLLDLASHYRVVPWNA
HGPLVGQTLAGSKIRQRYRINVLGVRPHTNKRPGDKPRLEAPTPDYVIRDGDTLLLVGDSDDV
SRFVAEVGG

SEQ ID No 104 (>ORF5)

SGSSGGGSSAEGSRCQPSGGGPHWLLEGETVTFPVTCASGLALAGDAFEVGPLPEGAAYDPIA
REVTFSPGLDQAAVYDIEIRVAQTSEVGRVKVGVADAFADPSNVPVVDPTRYPEEYGLPVLFL
SPVPEDKEYAPATVIYRGHTYAAEAELRGESSLSYPKRSYTLKFPKDDKFNEPDEAGGFTDRR
KVVLITTFDDNSYVRQRLAYDLWNRLDPEHIQIKTYSAVLYLDGEYAGLYTVADHVDGYLMED
HGYPQDGNLYKAVSHDANFALTDRSGDPKDTLHDGFEKKEGAPAEGEPEAFSDLEDLVSFVAE
SDDATFAAEIGSRIDLRDYEDWWIFVTFIVANDSAGKNSYHYRDPAADGVFRYAPWDFNASFG
QSWETEREPASDRVDYRDVNLLFERLLEEPSIGDPLRARYDQVLRGALAEAEIHAIVDGYVER
IDASARRDEARWGEAYRSYEGWSWRDDFTTYEEEIAYLK
AWISERWQHQDELY

SEQ ID No 105 (Contig 11 >ORF1)

VLDVWSTSDQVACRLHCAGAGPSASLELRYDASAGARRDAERLAERLAALLEDLSRHPERPVA
QGEVGPGERAEIEAWSRGPAMELPSACALHRWFEERAEQHPDVVAVRSEGKSLTYGELERRAN
RLASCLRRRGVGLDTIVGVCVPRSEDMVVATLAVLKVGGAYLPLDHEYPGERLAFMMRDARAR
LLVTHDAIADELPTGGWTTLLLDAEAAEIAACSDARPAVSPPPDSGAYVIYTSGSTGTPKGSL
ISHRAIVNQMQWIQRYWALTADDRVLLKAAFGFDVSVWEIFWPLSFGARIVVARAGGHRDPEY
LRRLVRDEGATTAYFVSSMLAAFLGGPEQPFPASLRKVLVGGEAVPLDLVRRFYAKHDGDLIN
MYGPSEAAIAVTGCVLPSDPRVTWVPLGAPVANAEVFVLDGAMRRPAIGALGDLYIAGAPLAR
GYVGQPGLTAERFLPDPCARAAGGRMYRTGDVARFLPDGMLEFQGRSDHQIKLRGHRIELGDV
EAQIRRVPGVGQAAVVLREDAPGDARLVAYVVLDGDAAGDAPDVRAGLKASLSAYMIPSSVVR
LYALPMCSERLAFTGSSYAGCLL

SEQ ID No 106 (Contig 11 >ORF2)

MSDHEMTGFSLSPQQRAIRALDREAGAPGCRTLAVVAVTGPCDEGRLSAAALALAERHEILRT
RLVEGRARPRRWSASRASRGRQQDDWVGCSEAEQGERMSRLVARLSEDRGADDGLRVGLVRVG
PEERRLVLAAPAWCVDEESIAPLVRELCASTAGAGAPPEQQYADVAEWLNGMLESEDAGDGRR
FWAERRSHFGPPLHLAFSRGGAGAGAGSGRARVDLQGGMAQVERWSSSWQVPQRIVLLALWAS
LLWRMSGGNEPEVTVAVRFDGRSLDALAGAVGPFARFLPVRIEISASDTLADVARRLALAEAE
AAAHQDAAPGVSHRMSWGLLRRGGRAGAVARRRAGPRARRLEHV

SEQ ID No 107 (Contig 11 >ORF3)

MSRIRAQLGVELPLRALFQGPTVAALAAQVDAARRGEARRREFPPIARIPRDGPLPLSFAQHR
LWFVDQLEPGSPAYNIPFVVRATGRLDVDALRRSLFEIARRHEALRTTFSARDGVPFPVVAPE
ARVPFRMSDLEHLAGEALDAAVSALVLEESLAPFDLSRGPLLRVRVIRKRHDEHVIALVVHHV
VFDVWSVGVFVGELAALYGGFAQGQPSRLPELPAQYVDFAAAQRAWLSGEVLEGELRYWTTKL
SGALRRARVPVDHEPAGRRTWRGARRSLDAGAELTRQIKAFCEREAISPFMALLAAYKLVLHQ
RTGLEDLVVGTDVANRNRVETEPMIGFFVNQLVLRTDCGGDPTFGALVRRVRDVALEAFEHQD
LPFDRLVEALRPKGAVGHVPLFDAKFVMRNVHVPPMKLEGLELEALEGEATTTAFDFVLTVAE

AGGSFRFGVEHSSELYRAATVDNFLSDYRQILATATARPDTPVSELRGELERAAAARRELERK
AARGAALDKLTSARRRAVTLPRPGAPGEAKTSPKDDLDE

SEQ ID No 108 (Contig 11 >ORF5)
MSEPIETEDGGSDIAIVGMAGRFPGAPSVDALWENVRRGVESIARFPESEREEPPVGASAAPG
APVVCAGGLLDDIDRFDASYFGYSPREAQLMDPQQRLFLECAVAALEDAGCDPARFPGAIGVF
GGCGSNTYLLQLLSHPDLAATVDPHALMLASEKDYLATRVSYKLDLHGPSVVVQTACSTSLVA
VHMACESLLGGQCDLALAGGVSIGIPQKRGYPYVPGSICSPDGRCRPFDARAEGTVGGSGVGI
VALKRLADALRDRNTVHAVIRGSAVNNDGGRKVGFMAPSVDGQAAAISEAQSVAGVDPGSIGY
VEAHGTATAIGDPIEVEALTQAFRRKTPRKAYCALGSIKANIGHLDAAAGVAGLIKAAHVVRS
GEIPPCVHFEAPNPKLDLAASPFFVPREAAPWPRELRPRRAGVSSFGIGGTNAHVVLEEPPPL
PPRAPAPERDHVLTLSARTPEALSTACAQLAAHLEATDVPLDDVAFTLQTGRAEHPYRRAVVA
RTRAEAIQGLAREGASALARPDEPRPSSRSRARARRPSGWPARSTRRRRSGAPSTRARRRRGR
AASISARSSSARARATGARCSAPRWRSPRSSPSSSRSPGSG

SEQ ID No 109 (Contig 11 >ORF6)
VVDHHVVVEYWSFALIVRELGELYSALRAGRPPQLPPPSSFFAAGVSCPSPREAAGGAEYWRK
ALDGTTAIDLPRDRARHDAGARRGRAHAITLPKPLTGALARLARERGTTLFSVLLSALTVLLH
RASGQSDLVVGVPSAGRHDDESARAFGYFVQMLPVRVALRGAASFDALVARVRDAFLDALAHG
DSALRHLLARAQGEAQRDALFDVAFAFQSTPPSLDARSALAIGVGDVRIAQGELELTTLADEQ
AAAEFDLALFAAELDAGIALRFEYDQQLFDPATIERMARHFVVLLESAVEHPGRPLSELRMLS
DAERALLLDDWSGAAAARQAASAPAPACVHALFEAHAARQPDATALEFGHQRFTYAQLSTWST
ELALWLRDRGVGPGSVVGVCIERSPRMVAAQLAVLKAGAAYASLDPANPPARLAEMLADCRAS
LALTSSQASHKLTAAPCPVHLVQDGACAPSTHIPLVSRPDDLAYVLFTSGSTGTPKGVCVRHA
SLSRLVSFLHLRLDLSPSDRWTQVASSGFDASVYEIWTPLACGAALLLADDDALRSPTALVSW
LVAQRATLSFMPTPLAEACFEQDWTGSALRAMTVGGDKLHPLRRPPFRLFNMYGPTEATVITT
VAEIADLGAEPPLGRPVDSALVYVLDPHMQPVPPGALGELYIGGACLAQGYTRTDLTAERFLP
DPFGQPGARLYRTGDLVRWRPDGQLAFAGRRDEQVKLRGRRVELGEVESVLRRLPGVREGIVV
LHGQGSAAHLIAHVVPDAHPPSERDLREGMARLVPDALVPAHFVLLPALPMSLSGKVDKKLLP
APPAAHADYEPPSGELELELAHIWQSVLHLDRVGRHDSFFDLGGHSLLAMQVLGRIESSLGIR

TTLRTLFEHPTLAQLASHLSSGAASTSAAAATALERGLTRPDGPSSPRVATPEEPFALTEGQR
AMWLECQKSADG
ALYNLGRTVRLGAGVDVAALRRAFEGLVERHEALRTTFLTRDGHPLQQVHRHVALEWAEEPAM
ALDEREIVARADEVRRRAFDLERGPLLRVHVWRRGEGQPPLLTVVVHHLVVDYWSFALLVREL
GELYSALRAGRPPQLPPPSSFFAAGVSCPSPREAAGGAEYWRKALDGATTAIDLPRDRARHDA
SPRRGRAHAITLPKPLTGALARLARERGTTLFSVLLSALTVLLHRASGQNDLVVGVPSAGRND
DESTRAFGYFVQMLPVRVALRGAASFDALVARVRDAFLDGLAHGDSALQHLLAEPRGAARRGG
ALFDVAFAFQGALPSLDPRLAALTTGAEDVRIAQGELELTTLADEQAAAEFDLALFAAELDSG
IALRFEYDQQLFDPATIERMARHFVLLLESAVEHPGRPLSELRMLSDAERALLLDDWSGAAAA
RQAASAPAPACVHALFEAHAARQPDATALEFGHQRFTYAELSTWSTELALWLRDRGVGPGSVV
GVCIERSPRMVAAQLAVLKAGAAYASLDPANPPARLAEMLADCRAALVLTSSQASHKLTAAPC
PVHLVQDGACAPSTHIPLVSRPDDLAYVLFTSGSSGTPKGVCVRHASLSRLVSFFQHLLALSP
RDRWTQLASSGFDASVYEIWTPLACGAALLLADDDALRSPTALVSWLVAQRATLSFMPTPLAE
ACFEQDWTGIALRAMTVGGDKLHPLRRPLPFRLFNMYGPTEATVITTVAEVADLGDEPPLGRP
IDSALVYVLDPHMQPVPPGVLGELYIGGACLAQGYTRTDLTAERFLPDPFGQPGARLYRTGDL
VRWRPDGQLAFAGRRDEQVKLRGRRVELGEVESALRRLPAVREGVVVLHGQGSAARLIAYVVP
GADPPSERDLREGMARLVPDALVPAHFVLLPALPMSLSGKVDKKLLPAPPAAHADYEPPSGEL
ERELAHIWQSVLHLDRVGRHDSFFDLGGHSLLAMQVLGRIESSLGIRTTLRTLFEHPTLHQLA
DRLSSGAASTTAAAATVPASEIAPSLGRAPAD
EPYPLSYEQERLWVLEQLLPGGTAYNVVQAVRLRNLVDVDALSSALAALVRRHWSLRTVFVAS
PTPQKICEPEAAPAEVVDLRGTPPDEAEAAARAWASREQATGFDLARGPVFRARLFRLDHDVC
VLVLSTHHIVTDAWSFQPLVRDLAELYRRARGGGPADMPELPLQYVDFAVWQRRHLAGKRLAD
KLAHWTATLRGLPVLELQTDRPRPPVQTFRGAERVLPLDARLVAQLDELARSRGATRFMVLLA
ALGVLLRRSSGQDDLAIGTAVANRPRPELEPLVGFFVNTIVMRLDLGGDPTFEELLSRARKVA
LEAFEHQDAPFEKVVEAVNPRRDLSRSPLFQVMLVVQNAPTEALELGEVRIEPLDLPVEATRF
DLRFSVEPRGGRDVISLQYNVDLFDAATIDRMLATMQSVLSRATQDPAQRVRALSVAPEDRER
ALVAWNDTAVATPDHLRLEEPFFERAVEQPDACAVVDAERRLTYGELARRAEAIAAAASRSGA
TANALVAVVMEKGWEQVAAVLGVLRAGAAYLPLDPRLPEERLRHLLEHAEVRLVLTQSAVDGT
IAWPAGIERLAVDADERWREQPVARRPPGGSTDDLAYVIYTSGSTGLPKGVMIDHRGAVNTVL
DINRRFDVGPEDRVLALSSLSFDLSVYDVFGTLAAGGAVVIPDRTRASDPGHWRELVERERVT

VWNSVPALMEMLMDASPGAGDPALSSLRLVMMSGDWIPLKLPDRIRAACRAPRVVSLGGATEA
SIWSIAHPIADVDPAWRSIPYGRPLANQHTYVLDEGLEPCPIGVPGEIHIGGIGVALGYWRDE
ARTRERFLKHPTTGERLYRTGDLGRYFADGTIELLGRTDHQVKIRGFRIELGEIEAALAQHPS
VEQAVVAAKTDPSGEKRLVAYVVGADGDGAALRDFVRKKLPEYMIPAEVVVLPALPLSANGKV
DRAALPDPAAVAPRAAAVAPRTATERLIASVLAEVLQVEAVGVTDNLFELGFTSLLLVRAQRL
LAERIAARAPDEGAAAQAVSLTDLFQYPTIEQLAQRLDAATVKAEPADVGAQRAEARRDARRR
RGRG

SEQ ID No 110 (Contig 12 >ORF1)
PPAVRRYVADRRPEQLPALAPEEREAAARRLSALGAAPPQVRRRGLTRAPLSYGQSRIYFLEQ
LSPGKPLFNVPGAVRLRGPVDVARLSAAFGEIVRRHDALRTSIANVDGELLQIAQPHAGFALD
VVTSTPEEAAELDRRLRAEAWRPFAIGAPPLLRATLFRLAEDEHVLLVTMHHVVSDDWSLGVI
LRELLALYAGRSLPPPRLQVSDFAAWQREMVESGALDGQRAYWRERLRGLSRASISAGGGAEA
PSHDPSGAIEEIALSPDKAAALEALARREGATLFMVLLALLDLVIHARSGALDIAVGTPIANR
NRPELEDVVGLLTNTLVIRVDLARAGAFRDVLARARVQALDAFANQDIPFDVVTQDLKQERDH
AQHPLFRVWLALQNAPKPALEVRGLRVEPLPLRPELVHFEVALLLWPADDGSVVGHFEFRRDR
VDEGARKEIAAAFTHLVDAVIARPDAPVSTLVEGARAEAARAQAALGEAFARAATARLGQLRR
RSAGDRTPRE

SEQ ID No 111 (Contig 12 >ORF2)
MSIHIEEEGRADAAKAPPFDYLQALHSALAHENDPVKRKQIEAGMVFKWLREEPLPFLSQLRR
EKPIFSIPAITLVTRYNDVVEVLNANDVFSVDNIAPKLVENVGQNILAMNDSPKYEHEKSLLR
LAFPRADLPRYRQIVVDEANRLLAKVGVDKPFDLTGDYALRVPAGAMARYLGVGEIPTEKVVA
WTHALFNEIFLNPTNDPTAVAAARAARQEALPMIDAIVAARKKQLAKSPPPEQPSVLDRYLVM
QSVPETYESDEGIRDVILGLLMGCVDLSGGAIVNALVELMKRPRVLRDALNVVNVEDDAAITG
YVLEALRFRPPSTGVTSLCVRDYTVGRGTRHEEKVPAGALVMACSASAMHDHEHIDAPDQFRP
GRLPSRNYLFWESGIHTCHGKYVAILHISLAIKQLLRAGVPSAIDPMPRVHGYPAPFRVRLAA
AEG

SEQ ID No 112 (Contig 12 >ORF3)

MREPSSTPNWRNFGSNLPAGSDSVPPGEGFPIKKILALNLGKWKDTAGLQIAQALHLFEYGYK
RYREGKFVLRATSDLGLGAIFESIDNWESFDQFEEFFKPWTFIRKPLVATRWAEDAEFGRQRL
VGINPAHIRRATPADLADFVSGAEPKPIAIADGRTLEEVREGGQLYFLDYRIFKDIVDTDVQE
ELGKYPLAPTCMLHQTAAGELLPVAIRLVHSRPGKGAHPDKIFTPSGPSDDWLTAKIAVASAD
AIYQGQVTHLLYAHLIVEPFAVSTYRNLPATHPLHQLLRPHFFNTLAINELARRRFLGRGRFF
DITSSVATMGSFELLTRAYTGKGIKGYGGKPWRFYESALPRDLSARDVRDLVGYHYRDDALLH
WDAIQEYVGQVLKIAYPTPGSLSSDASLQRWIHELVSPQLGGMDSLLPPERADQLEKLTSLDD
LIAIVTNIIFTATAYHAAVNFGQTDYYTWIPNAQFATYRSYGDVLNGSEKRQFKPLERLPGRA
QSIRQMVLSRSLSMGPPLTSESLMTMKCLLQDPAAKQAFARYRERLAHIEREITERNRAREQP
YLYLLPSMVPQSVAI

SEQ ID No 113 (Contig 12 >ORF4)

VSSSRSTGRVPRDRASPAGSCAPALVPGPPLSYASVMPPLDLHVALFGASGAGKTVLLAAFYR
AQTQPSFQQEYAYKIQAVNKAQGNQLLGRFYRLEEGRFPDGSTRFDEYEFDFFPRDLPEPAVR
IHWYDYPGRWWEDEPVDADEREAMRQGLIRLGMSQVGILLADGAKYRAEGTGYIRWLFEHFAD
ECDRLRRASAATGDEVSFPREWILALSKADLCPPDYSARDFEREVCRDADDQLAKLCSVLRAE
HAFGHRFMLLSSVAAPAGAQVDPRTSLGVRTLAPAILVSTVEGAVREAQAARKEKSAGETFFQ
GLRDLVQFVDSLDDFLPKRYQIVSKILRFISIKDFATTRLDRLKKMREDAIRKGDTFTAVLTA
MVAALRDDEGARAYHQNQ

SEQ ID No 114 (Contig 12 >ORF5)

MPAPAPLVETSRLLWRTRGEHWDYEFICVPEIPALPAWLSTLEAMLADADAGAGELRYGLLEI
DDRGQRAPRAYPYVAVRFLDPARRDWTGRQVQHFAAWFPPVPPEAVAELPEAVPADWHLRVLD
GLAGTYGSGEVFGLPEATIRAWKRSHDESRAARAMAIVKATPPVSLGGGEAAPSRWTRVPTLK
KKPPEPPAAAGLLSVGAVPSGQGRRFGCFAIGAMMLAAFCRLMLACGVRLLGA

SEQ ID No 115 (Contig 12 >ORF6)

VRFRSSLGPLLLAALGAALTVSAAPRSAEASVFDSASRWPEDADGHVRIPVCIDPTSSAEQRV
DGAAGGLIHAPNPSLADVITRVRTALQGSWERWSSVRFTGWESCDSLLPATRMTYVGVRIHPD

APNQSDSIGVYNKGGSVQFKPWGADFNRCIKYNWQTARVEYSFDCVEQYAIHEMGHAIGFMHE
WHHPLVPSACSQREPLPASDVASGWPSSRRYIVVNPGFYDYDSIMTYWSGCSDQDGVRFGSET
LDAVDIQAVATVYPPVGGAPDVCNPGWFAGKRWFCAAQPTVSVGNSCSSGWVECLPHCNPRPF
QGEWWTCPTNPYAVTGQSCSARWELCGD

SEQ ID No 116 (Contig 12 >ORF7)
VGESQGALVGGNALSTNALNLNALNLNALNLNALNLSGLSARNLAAIQDPGPSGALARDFLRY
AASCALSSTASFDFSWTDSNGKRHDERYPGLLGVAPAWASGPLDDAGQRIVSSCVAARVNYYQ
VPVLLSARSLRDPLKTLSSSQELIDYPDVEGAFWGNLFAAQPYINACYNSATVDNSRAYQRDC
AAGHVTSGGQIVECGLIRIAGSCDRVCQKLNGAGQYYPSCVDRPGQSTATTKDVITTALP

SEQ ID No 117 (Contig 12 >ORF8)
VLAAHCERGGLTARAASLLARGAELAAARRAYLDAEGCYGRVEALLGALLPEERRARGLARFR
LGRHTEALADLAAAREAAAAASEAGAEIELLLDEAMILDWTGEYRAARERVAAAERLAGRVAS
PLLGARLLLGVGRSLHRADREDEAAAVLTRAAAQAARLGDEGHETHIIALLLLGFILASLGRV
EEAARDLDAVILSCEERSDLMHLGAALNNRGLARALQGDRAGMIADFERTIALGRELGQPAFE
LVGRYNLAEYLYLMDDLAAARPHARAVQAIAPRCGDRHAPVVVTLLIARLRLYQGDEAGARRI
ALRLRAARDDAGCEALKPSEDVLCAMIELATRDDDRAAWAALEERSARCSVGQERIEVLEARA
LAALRRGRRADARAQLERALAAASTIPTVMGGRLRRWYAELTRATESDAPDIDLAAAEATFTG
ARAREKVEY

SEQ ID No 118 (Contig 12 >ORF9)
QAYPDLWAERGRQELWLRQLPPRACAQLAREALGDAADGALIDRLVTQSEGQPFFLEELIRAT
AEGRGDALPETVVAMVQVRLEALAPPARRILRAASVLGEVFWRGAVAHLLGGDEAAPLAEHLS
ALVAGELCVRHREGRFPGEEEYSFRQALLREGAYAQLTKDDRALGHRLAADWLEAAGEADPLV
LAAHCERGGLTARAASLLARGAELAAARRAYLDAEGCYGRVEALLGALLPEERRARGLARFRL
GRHTEALADLAAAREAAAAASEAGAEIELLLDEAMILDWTGEYRAARERVAAAERLAGRVASP
LLGARLLLGVGRSLHRADREDEAAAVLTRAAAQAARLGDEGHETHIIALLLLGFILASLGRVE
EAARDLDAVILSCEERSDLMHLGAALNNRGLARALQGDRAGMIADFERTIALGRELGQPAFEL
VGRYNLAEYLYLMDDLAAARPHARAVQAIAPRCGDRHAPVVVTLLIARLRLYQGDEAGARRIA

LRLRAARDDAGCEALKPSEDVLCAMIELATRDDDRAAWAALEERSARCSVGQERIEVLEARAL
AALRRGRRADARAQLERALAAASTIPTVMGGRLRRWYAELTRATESDAPDIDLAAAEATFTGA
RAREKVEY

14. DNA sequence according to any of claims 1 to 5 wherein the DNA is selected from the group consisting of

(a) the following DNA sequences:

Seq ID No 119 (>Contig17)

TTACGTTACTCATCCTATCTCGGCACCCTGTGTCGGTGATGTCGCTCGCC
TCGAGCGCGAGCGGGACGACGTCGGCGCCGCGCTCGGTGAGCGCCGCCGC
GAGGGCGCTCGCGAGATCGCTGGCGACGCCGGCCGGGGCCACGACGAGCC
ACGTCCCCGCGACGTCGCCGCGTGACGCGGCGCTCACGGGTCTCCATTCG
ACGCGGTAGCGCCACGCGCCCACGGTGCTCTGCTCTCGGCGGCTCCGCCG
CCACGCCGACAGGGCCGGCATGAGGCTCTCGAGGGCCGAGCGCCGCCCGC
TGTCGGCGACGTGGAGCGCGTCCGAGAGCGCCGCGACGTCGCCGCGCTCG
ATGGCTCGCCAGAACGCGGTCTCCTCGGCGGACGCTCCCGGCGCCGCGTC
CTCATCGTCCGACGCGTCGCCTGCGTCGAGCCAGAACCGCTCGCGCTGGA
ACGCGTACGTCGGCAACGTCACGCGGCGCGCCCCGAGCGGAGCGAAGAAC
GCACCCCAGTCGATGGCGTGCCCGCGCGCGTGGAGCTCGCCTGCCGAGAG
GAGGAAGCGCTCGAGGTCGCCTTCGTCGCGGCGGAGCGAGGACACCACGG
TCGCATCGCCGTCGATCGACGAGAGCGTCTCGTCGAGCGCGACGGTGAGC
ACGGGGTGAGGGCTGACCTCGACGAAGAAGCGGTGGCCGTCGTCGAGCAG
GGCGCGCGTGGCGTGCTCGAAGCGGACGGTGTGGCGCAGGTTTCGGTACC
AGTGGGCGGCGCCGAGGGCCTCGCCATCAAGCCTCTCGCCCGTCACCGCG
GAGTAGAGCGGCACGGTCGCCGGGCGCGGCGCGATGCCGTCGAGCGCCTC
CAGCATCGTCCGCTCGATGGCCTCCACGTGGGCGGAGTGGGAGGCGTACT
CGACGCGGACCTTGCGGGCGAACAGCTGCGCCCCGCTCAGCTCTGCGACG
AGCTCGTCGATAGCGCCGGGGTCTCCGGAGACGAGGGCCGCGTGAGGGCT

GTTGATCGCCGCTATCGCCAGGCGTTCGCCCAAGGGCGCAAGGCGCGCCT
CGAGCTCGGCGGTGGTGAGCTCGACGGCGGACATGGCGCCGCGTCCCGCG
AGCTTCGTAATGGCGCGCGAGCGGAGCGCGACGACCCTGGCGGCGTCTTC
TAGCGAGAGCGCGCCCGCGACGTACGCGGCCGCGATCTCGCCCTGGCTGT
GGCCGACGACCGCGTCGGGCGTGACTCCGGCGGCGCGCCAGGTGGCGGCG
AGGGCGATCATGACGGCGAACAGCACGGGCTGCACCACGTCGACGCGCTC
GAGCATGGGCGCGGCGTGCGCTTCGTCGCCGCCGAGCACGGCGAGGAGCG
ACCAGTCGACGTGCGGCGCCAGGGCGCGCTCGCACGCCTCGATCTCGGCC
CGAAAGGCGGGCGAGGAGGCGAGCAGAGCGCGCGCCATCGATGGCCACTG
CGAGCCCTGGCCGGGGAAGACGAAGGCGACCTTGCCCGGCGGGAGCGCCT
CGCCCGCGACCGTTCCTGCCCCCGCGCGCCCCTCGGCGAGCGCCGCGAGC
GCCGAGAGCAGCGCGGCGCGATCGTCTGCCACGACGGCGGCGCGACGCTC
GAAATGCGACCGCGTGGTCGCGAGCGACGCCGCGACGTCGACGAGGGCGA
CGTCCTCGTGCTCGGCGAGGTGCGCGTGGAGCTTGCCCGCCTGAGCGCGG
AGCGCCGCGTCGCTCTTCGCCGAGAGGAGCACCGGCACCGGCGGCGCGAA
GGGCGCGCGGGCGGGCTCCCCGGCCTGGTCGTCGCCGGCCGCCGCGCGCG
GCGCTTCCTCGAGGACCACGTGCGCGTTGGTGCCGGAGATCCCGAACGAC
GACACCGCCGCGCGCCGAGGAGACCCGCCTGGCTTCCACGGTACCTCCTC
GGTCAAGAGGCGGATCGCGCCGGACGACCAATCGATGTGCTGCGACGGGC
TCGCGGCGTGGAGCGTCCTCGGGAGGACGCCGCTCTGCAGCGCGAGCACC
ATCTTGATGACGCCGCCGATCCCCGCGGCGGCCTGCGTGTGCCCGAGGTT
CGACTTTAGGCTCCCGAGCCACAGCGGGCGCTCCTTCGCGTGCGCCGCGC
CGTACGTCGCGAAGAGCGCGCGCGCCTCGATGGGATCGCCGAGCGTCGTG
CCGGTTCCGTGCGCCTCGACGGCGTCGACGTCCGCGGGGGCGAGCCCCGC
GCTCGCGAGCGCGTCCCGGATCACGCGCTCTTGCGCGGGGCCGTTCGGCG
CCGTGAGCCCTTGGCTCTTGCCGTCCTGGTTGACGGCCGATCCGCGCACG
ATCGCGAGCACGGGGTGCCCGTTCTTCCGGGCGTCCGACAGGCGCTCGAG
GAGCACTATCCCAGCGCCTTCCGACCAGCCCGCGCCGTTCGCGTGCGACG
AGAACGACTTGCACCGCCCGTCCGGCGCGCCCGCGTGCTGCGCGCTGAAC
TCGCCGAAGATCCCGGGGGTCGCCATCACGGTCACGCCGCCGGCGAGCGC

GAGCGAGCACTCGCCTCGACGGATGGCGTGGCAGGCGAGGTGGAGCGCGA
CGAGCGACGAGCTGCACGCCGTGTCGACGCT

Seq ID No 120 (>Contig18)

TTTTAGGANCCCCGACGTGCACGATCGGCTCGCCAACCTCGTGGCGCGCC
GGGACTATTTTTACCAGCTCGCGTTGCGCGCCGCGGGGACCTACGTGCGG
GGCCTCGTCCGCGCCCCGCACGACGGCGCGCGCCCCCCCGCGTTCGCGCC
GCGTGGGGCGGCGCTCGTCACGGGCGGGACCGGGGCGCTCGGGGCGCACG
TTGCCCGTTGGTTCGCGCGGATCGGCGCCGAGCACATCGTGCTCGCGAGC
CGCCGCGGAGCCGCGGCCCCCGGCGCGGCCGCGCTCGCCGAGGAGCTTTC
GGTGCTCGGCGCGCGCGTGACGCTGGTTGCGTGCGACGTCCCCGATCGTG
AGGCGGTCGCGGGGCTCGTGCGCAACGTCAAGGCCGGCGGAGCGACGGTG
CGCGCCGTGTTCCACGCGGGCGGTGCGATGCACGAGGCGCCGGTCGCCGC
CATGCGTGTTGAGGAGCTCGCCGACGCGATCGCCGTGAAGGCCCGCGGCG
CGCAGCACCTCCAAGACGTCTTCGCGCAGCGCCCGCTCAACGCGTTTGTC
CTCTTCTCGTCAGAAACGGGTGTGTGGGGCGGTGGCCGGCAAGGCGCGTA
CGCCGCGGCGAACGCGTTCCTCGACGCGCTCGCCGAGGCGCGTCGCGCGG
ACGGCCTCGCGGCGACCTCGATCGCGTGGGCGCGTGGGCGGGCGGCGGA
ATGCTCGCGACCGACGCCGAGCGGCGCTTGAAGCATCGCGGCGTCGCGCC
GATGGATCCGGAGCTCGCCGTCGCGGCCCTCGCGCACGCGCTCGATCACG
CCGAGACGTGCCTCGCCGTCGCTGACGTCGACTGGGCGCGCTTCGCCCCG
TCGTTCGCCTCGGCGCGTCCTCGCCCGCTCCTCGACGAGCTCGCGGAGGC
GCGATCGGCGCTCGACGCGCTGCGCGAGCCACCGGACGACGCGCGCACGG
CCGCCGGTCCCGAGCCCGCAAGCACGCTGAGGACCACGCTCGCGGCGCTC
CCGGAGGGCGAGCGCCACCGCCACCTCCTCGCGCTCGTGCGGACGGAGAC
GGCGGCGGTGCTCGGGCACGCGGACGCGTCGCGCGTCGAGCCGAACCGCG
GGTTCTTTGACCTCGGGCTCGACTCGCTCATGTCCGTCGAGCTCCGCAGG
CGCGTCCAGCGCGCGACCGGCATCAAGCTCCCGGCGACGCTCGCGTTCGA
CCACCCGACGCCGAGCGCGCTCGCGAGCAAGGTGCTCGCCGCGATCGTCC
TCCACGACGCGACCCCGCGCGCCTCGCCCGCCGCGGAGCTCGAGCGCCTC

GAGGGGATGCTCTCGGCGATCTACGCGGACGAAGCGCTCCGCGACGACCT
CACGGCGCGCCTCCGCGCCTTCCTGGACAAGCGCGCGGTCCGCACCGAAC
GCCCCGACGACGCCGCGTTCGCCGAGAAGCTCGGCTCCGCGAGCGCCGAC
GAACTCATTCGCCTGATCGATCAGAAGCTCGGAGATCGCATCGATGTCGA
CCGTTACTAACGACACGCTCACGGAGTACTTGCGGCGCCTCACTCAAGAG
CTCCACAGGAGCGAGACGCGCCTGCGTGCGACGGAAGAGAGGCGACATGA
GCCGATCGCCATCGTCGGCCTCGGGCTCCCCTTCCGGGGCGGGATCCACG
ACCGCGACACGCTCTGGACGTTCCTCGAGGAGGGCCGCGACGCCATCGCG
CCGATCCTCGCGAGCCGCTGGAACGCGGACGCGACGTACGACCTCGATCC
GGACGCCGTCGGCAAGAGCTACGTGCGCGACGCCGCCATGCTCGATCGCG
TCGACCTTTTCGACGCCGATTTCTTCGGGATCAGCCCGCGCGAGGCGAAG
TACGTCGACCCGCAGCACCGCCTCTTGCTCGAGACGTCGTGGCAAGCGCT
CGAGGACGCGGGGATTGTGCCGGCGTCGCTGCGAGACTCGAAGACCGGCG
TCTTCGTCGGCACGGGCGCGAGCGACTACGCGTTCCTCCAGAGCGATCGC
GACGCCTCGGAGGCGTACGCGTTCATGGGGATGATCTCGTCGTTCGCGGC
GGGCCGCCTCGCGTTCACGCTCGGGCTCCAAGGCCCCGCGCTATCGATCG
ACACGGCGTGCTCTTCGTCGCTCGTCGCGCTCCACCTCGCGTGCCAGTCG
CTGCGTCAAGGCGAGTGCGACCTCGCGCTCGTCGCGGGTGTGCAGGTCAT
GTCGTCGCCGGAGGTGTTCGTGCTGCTCTCGCGCACGCGCGCGCTCGCGA
GCGACGGGCGATCGAAGACGTTCTCGGCGAACGCCGACGGCTATGGCCGC
GGCGAAGGCGTCGTCGTCCTGGCCGTCGAGCGCCTCCGCGACGCGCGCGC
GAAAGGGCGCCCGATCCTCGCGGTGATCCGCGGCAGCGCGGTGAACCACG
ACGGCACGTCGAGCGGGATCACGGTCCCGAACGGGCCCGCGCAGCAGAAG
GTGCTCCGCGCCGCGCTCGACGACGCGCGGCTTGTCCCCGCCGACGTCGA
CGTCGTCGAGTGCCACGGCACGGGGACCTCCATCGGCGATCCCATCGAAG
TGAACGCGCTCGCCGCCGTCTACGGCGAGGGGCGCCCCAAGGACCGCCCG
CTGTTCCTGGGCGCGCTGAAGACCAACATCGGGCACCTCGAGTTCGCGTC
GGGCCTCGCCGGCGTCGCGAAGATGGTCGCCTCCATGCGCCACGCGACCC
TCCCCGCGACGCTGCACACGAGCCCGCTCAACCCGCTCGTCGACTGGGAC
GCGCTCCCCGTGCGCGTCGTCGACGCCGCGCGCCCGTGGACGCGCCGCGA

CGACGGCGCCCCCCGGCGCGCCGGCGTCACGGCGATCGTCGAGGAGGCGC
CCGCCGAGCCCGAGCCCACGACGCCCGACGCCGCGCCCGCGCTTCCGGCC
GTGCCCGTTCTCCTCTCGGGCAAGACCGACGAGGCGCTGCGCGCGCAGGC
AGCGCGCCTCCACGCGCACCTCGCGGGGCGCCCCGACGCGCGGCTCGTCG
ACATCGCCGCGTCGCTCGCGACGACGCGCACGCACTTCGATCGACGCGCG
GCCGTCGTCGCGGCGGATCGCGACGAGCTCCTCGGCGCGCTCGACGCGCT
CGCGCGCGGCGAGGCAGGCCCGGGGTCGGTCGTCGCGAGCGCGATCCCCG
CCGGCAGGGTCGTGTTCGTGTTCCCCGGCCAAGGCTCGCAGTGGGTCGGG
ATGGCGCGCGCGCTCCTCGCGTCGTCGGTGGTCTTCCGCGACGAGATCGC
GGCCTGCGAGCGCGCGCTCGCGCCGCACGTCGCCTGGTCGCTCGGCGCCG
TTCTCCGGGGCGACGGCGACGAGGCGACGCTCCTCGGCCGCGTCGACGTC
GTGCAGCCGGTCCTCTTCGCCGTCATGGTCGCCCTCGCCGCGCTCTGGCG
CTCGATCGGCGTCACGCCCGACGCCGTCGTCGGGCACAGCCAAGGCGAGA
TCGCCGCCGCCTACGTCGCCGGCGCCCTCTCGCTCGAAGACGCCGCCAAG
GTCGTCGCGCTGCGCGCACGAGCGCTCACGAAGATCGCGGGGCGCGGGGC
GATGGCCGCCGTCGAGCTCGGCGCACGCGACACCGAGGCGCGCCTCGCGC
CGTTCGGCGACGCCATCGCGATCGCGGCGATCAACAGCCCGCGCGCCACG
CTCGTCGCGGGCGACACGGACGCGATCGACGCGCTCGTCCGCGACCTCGA
GGCCGCGCAGATCTTCGCGCGGAAGGTGCGTGTCGACTACGCGTCGCACT
CGGCGCACGTCGAGGCGATCGAGCGCGAGCTCCTCGCGGATCTCGCGGGG
ATCGAACCGCGCGCGGGCGCTGTGCCGCTTTACTCCGCGGTGACGGGCGC
GAAGCTCGACGGGAACCGCCTCGACCCCGCGCATTGGTTCCGGAACCTGC
GCTCGACAAAAAACTTTGAGGACGCCACGCGCGCGCTCCACGACGACGGC
CGCCGGGTATCCTCATNATCNNGGGCGTNCAGAGGAGTCGGTATTNCCCC
CCCCCGCCTTNCCCG,

or their complementary strands,

(b) DNA-sequences which hybridise under stringent conditions to regions of DNA-sequences according to (a) encoding proteins or to fragments of said DNA-sequences,

(c) DNA-sequences which hybridise to the DNA-sequences according to (a) and (b) because of a degeneration of the genetic code,

(d) allele variations and mutants resulting by substitution, insertion or deletion of nucleotides or inversion of nucleotide segments of DNA-sequences according to (a) to (c), wherein the variations and mutants offer isofunctional expression products.

15. Peptide encoded by a DNA sequence according to claim 14 selected from the group consisting of

Seq ID No 121
>Contig17_001 828 amino acids MW=86259 D pI=5.60 numambig=1
MTVMATPGIFGEFSAQHAGAPDGRCKSFSSHANGAGWSEGAGIVLLERLSDARKNGHPVL
AIVRGSAVNQDGKSQGLTAPNGPAQERVIRDALASAGLAPADVDAVEAHGTGTTLGDPIE
ARALFATYGAAHAKERPLWLGSLKSNLGHTQAAAGIGGVIKMVLALQSGVLPRTLHAASP
SQHIDWSSGAIRLLTEEVPWKPGGSPRRAAVSSFGISGTNAHVVLEEAPRAAAGDDQAGE
PARAPFAPPVPVLLSAKSDAALRAQAGKLHAHLAEHEDVALVDVAASLATTRSHFERRAA
VVADDRAALLSALAALAEGRAGAGTVAGEALPPGKVAFVFPGQGSQWPSMARALLASSPA
FRAEIEACERALAPHVDWSLLAVLGGDEAHAAPMLERVDVVQPVLFAVMIALAATWRAAG
VTPDAVVGHSQGEIAAAYVAGALSLEDAARVVALRSRAITKLAGRGAMSAVELTTAELEA
RLAPLGERLAIAAINSPHAALVSGDPGAIDELVAELSGAQLFARKVRVEYASHSAHVEAI
ERTMLEALDGIAPRPATVPLYSAVTGERLDGEALGAAHWYRNLRHTVRFEHATRALLDDG
HRFFVEVSPHPVLTVALDETLSSIDGDATVVSSLRRDEGDLERFLLSAGELHARGHAIDW
GAFFAPLGARRVTLPTYAFQRERFWLDAGDASDDEDAAPGASAEETAFWRAIERGDVAAL

SDALHVADSGRRSALESLMPALSAWRRSRREQSTVGAWRYRVEWRPVSAASRGDVAGTWL
VVAPAGVASDLASALAAALTERGADVVPLALEASDITDTGCRDRMSNVX

Seq ID No 122
>Contig18_002 502 amino acids MW=53019 D pI=6.83 numambig=1
FRXPDVHDRLANLVARRDYFYQLALRAAGTYVRGLVRAPHDGARPPAFAPRGAALVTGGT
GALGAHVARWFARIGAEHIVLASRRGAAAPGAAALAEELSVLGARVTLVACDVPDREAVA
GLVRNVKAGGATVRAVFHAGGAMHEAPVAAMRVEELADAIAVKARGAQHLQDVFAQRPLN
AFVLFSSETGVWGGGRQGAYAAANAFLDALAEARRADGLAATSIAWGAWAGGGMLATDAE
RRLKHRGVAPMDPELAVAALAHALDHAETCLAVADVDWARFAPSFASARPRPLLDELAEA
RSALDALREPPDDARTAAGPEPASTLRTTLAALPEGERHRHLLALVRTETAAVLGHADAS
RVEPNRGFFDLGLDSLMSVELRRRVQRATGIKLPATLAFDHPTPSALASKVLAAIVLHDA
TPRASPAAELERLEGMLSAIYADEALRDDLTARLRAFLDKRAVRTERPDDAAFAEKLGSA
SADELIRLIDQKLGDRIDVDRY*

Seq ID No 123
>Contig18_010 840 amino acids MW=88062 D pI=5.74 numambig=6
MSTVTNDTLTEYLRRLTQELHRSETRLRATEERRHEPIAIVGLGLPFRGGIHDRDTLWTF
LEEGRDAIAPILASRWNADATYDLDPDAVGKSYVRDAAMLDRVDLFDADFFGISPREAKY
VDPQHRLLLETSWQALEDAGIVPASLRDSKTGVFVGTGASDYAFLQSDRDASEAYAFMGM
ISSFAAGRLAFTLGLQGPALSIDTACSSSLVALHLACQSLRQGECDLALVAGVQVMSSPE
VFVLLSRTRALASDGRSKTFSANADGYGRGEGVVVLAVERLRDARAKGRPILAVIRGSAV
NHDGTSSGITVPNGPAQQKVLRAALDDARLVPADVDVVECHGTGTSIGDPIEVNALAAVY
GEGRPKDRPLFLGALKTNIGHLEFASGLAGVAKMVASMRHATLPATLHTSPLNPLVDWDA
LPVRVVDAARPWTRRDDGAPRRAGVTAIVEEAPAEPEPTTPDAAPALPAVPVLLSGKTDE
ALRAQAARLHAHLAGRPDARLVDIAASLATTRTHFDRRAAVVAADRDELLGALDALARGE
AGPGSVVASAIPAGRVVFVFPGQGSQWVGMARALLASSVVFRDEIAACERALAPHVAWSL
GAVLRGDGDEATLLGRVDVVQPVLFAVMVALAALWRSIGVTPDAVVGHSQGEIAAAYVAG
ALSLEDAAKVVALRARALTKIAGRGAMAAVELGARDTEARLAPFGDAIAIAAINSPRATL
VAGDTDAIDALVRDLEAAQIFARKVRVDYASHSAHVEAIERELLADLAGIEPRAGAVPLY

SAVTGAKLDGNRLDPAHWFRNLRSTKNFEDATRALHDDGRRVSSXSXAXRGVGIXPPRLX
X

16. Recombinant expression vector which comprises a DNA-sequence according to any of claims 1 to 10, 12 and 14.

17. Procaryotic or eucaryotic cell which has been transfected or transformed with a DNA-sequence according to any of claims 1 to 10, 12 and 14 or with a recombinant expression vector according to claim 16.

18. Cell according to claim 17, wherein the cell is derived from myxobacteria.

19. Cell according to claim 17, wherein the cell is derived from a Sorangium strain.

20. Cell according to claim 17, wherein the cell is derived from Sorangium cellulosum.

21. Cell according to claim 17, wherein the cell is derived from a Streptomyces strain.

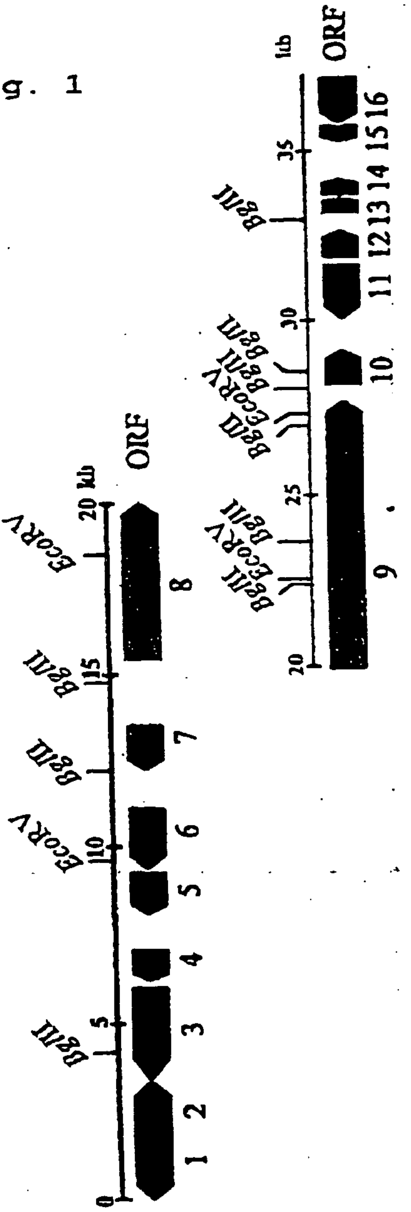
22. Cell according to claim 17, wherein the cell is derived from Escherichia coli.

23. Process for an enzymatic biosynthesis, mutasynthesis or partial synthesis of polyketide or heteropolyketide compounds, wherein a cell according to any of claims 17 to 22 is culti-

vated in a suitable culture medium and the polyketide or heteropolyketide compound is isolated from the medium.

24. Process according to claim 23, wherein the polyketide or heteropolyketide compound is an epothilone.

Fig. 1


0
5
10
15
20 kb
ORF
BglII
EcoRV
BglII
BglII
EcoRV
1
2
3
4
5
6
7
8
20
25
30
35
kb
ORF
BglII
EcoRV
BglII
BglII
EcoRV
BglII
BglII
BglII
9
10
11
12
13
14
15
16

Open reading frames found on pEPOcos6 region.

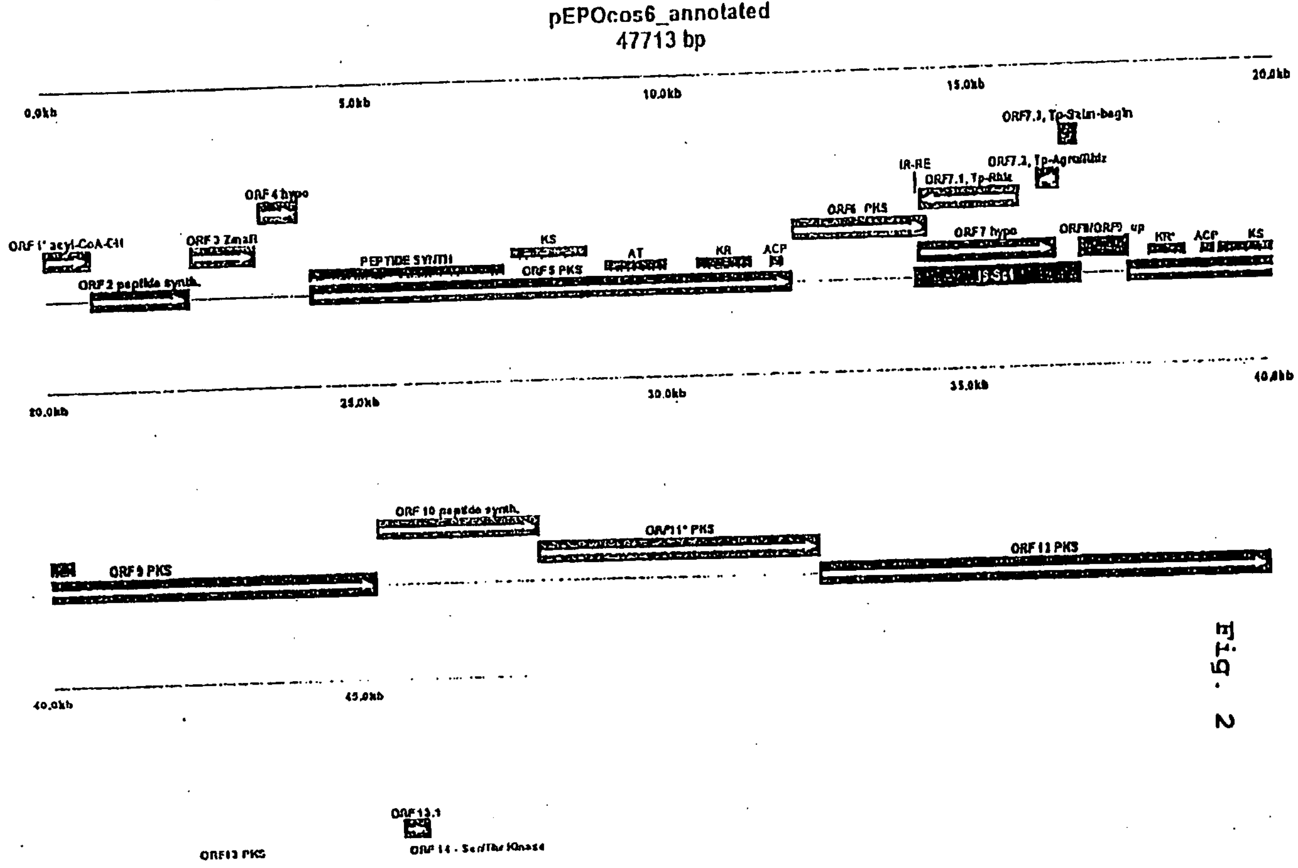


Fig. 2